Multi-User Non-Cooperative and Cooperative Systems with HARQ

Zahid Rauf

A thesis submitted for the degree of

Doctor of Philosophy

in

Electrical and Electronic Engineering

at the

University of Canterbury,

Christchurch, New Zealand

2013

نظر حیات پہ رکھتے ہیں مردِ دانش مند
حیات کیا ہے حضور و سُرور و نُور و وجود

On life is fixed the gaze of persons bright,

What is life? Presence, being, joy and light!

Dr. Muhammad Iqbal

(1877-1938)

Abstract

Multi-User Non-Cooperative and Cooperative Systems with HARQ

by

Zahid Rauf

The performance and reliability of wireless communication links can be improved by employing multiple antennas at both ends, thereby creating multiple-input multiple-output (MIMO) channels. However, once multiple co-channel users are added to the system it can be difficult to provide as many receive antennas as transmit antennas, resulting in a so-called overloaded (rank-deficient) system. Under overloaded conditions, maximum likelihood (ML) detection works well, but its exponential complexity prohibits its use and suboptimal linear detectors perform poorly.

In this thesis, new signal processing techniques for multi-user overloaded systems using hybrid automatic repeat request (HARQ) protocols are investigated. The HARQ retransmissions are used to form virtual receive antennas, which can efficiently transform an overloaded system into a critically loaded system (i.e. a system with an equal number of transmit and receive antennas).

In the first part of the thesis, a multi-user non-cooperative overloaded system is considered. At first, it is demonstrated that the suboptimal linear minimum mean square error (MMSE) detector leads to significant perfor-

mance degradation compared to an ML detector for such systems. To overcome this drawback, two multi-user transmission schemes are proposed that work well under overloaded conditions. The proposed schemes allow us to apply linear multi-user detection (MUD) algorithms without requiring additional antennas or hardware chains. Monte-Carlo simulations demonstrate that the proposed schemes can result in significant gains in terms of bit-error-rate (BER) and dropped packet performance.

In the second part, the performance of multiple HARQ processes for a two-hop multi-source multi-relay decode-and-forward (DF) relaying network with no direct link are analyzed. Dealing with multiple HARQ processes at each relay, a retransmission scheme is proposed that utilizes virtual antennas to achieve increased receive diversity and improved throughput compared to traditional orthogonal (time division) retransmissions. A novel forwarding strategy at the relay(s) to destination link is proposed with the objective of further improving throughput. Finally, the end-to-end outage probability and throughput efficiency of the proposed retransmission and forwarding schemes are found analytically and confirmed with Monte-Carlo simulations.

To my parents, Abdul Rauf and Noor Jan,

to my wife Rizwana

and to my kids, Bilal and Abyan.

Acknowledgements

This research work would not have been possible without the help and support of many people.

First of all, I would like to express my sincerest gratitude to my supervisors, Prof. Desmond P. Taylor and Assoc. Prof. Philippa A. Martin for their continuous support, guidance and encouragement over the years. Their constructive suggestions, insightful comments and endless proof readings helped me in completing this thesis. I am thankful to them for not only their academic guidance but also for the non-academic support, especially the financial assistance during the last year of my doctoral studies.

I would like to thank all my friends and colleagues at the communication research group (CRG), past and present, for their great support and useful suggestions during my stay at CRG, particularly Dr. Rui Lin and Tauseef Tasneem. I would also like to acknowledge my other friends living in New Zealand, who made my years so enjoyable.

My special thanks go to Dr. Samudragupta Bora (Sam) my flat mate and dearest friend for always being so supportive and caring. Thanks for the much needed discussions that always reduced my stress and kept me working. Life was wonderful being with you amidst all ups and downs.

I would like to acknowledge the higher education commission (HEC), Pakistan and Balochistan university of information technology, engineering and management sciences (BUITEMS), Quetta, Pakistan for their financial support during my doctoral studies.

I am sincerely grateful to my beloved parents for their everlasting love, prayers and encouragement. If they were not there for me this journey would have never been possible. I am also thankful to my brothers and sisters for their constant support and encouragement. I am deeply indebted to my wife, Rizwana, for her love, patience and support. Thank you for taking care of our boys, Bilal and Abyan, while I was away from home.

Most of all, I am thankful to Allah (God), the exalted for blessing me the ability and strength to complete this research work.

Contents

List of Figures

List of Tables

Acronyms

ACK	Positive Acknowledgment.
ACO	Ant Colony Optimization.
AF	Amplify-and-Forward.
AP	Access Point.
ARQ	Automatic Repeat Request.
BER	Bit-Error-Rate.
BS	Base Station.
CC	Chase Combining.
CCI	Co-Channel Interference.
CDF	Cumulative Density Function.
CDMA	Code Division Multiple Access.
CRC	Cyclic Redundancy Check.
CSI	Channel State Information.
DF	Decode-and-Forward.
e2e	End-to-End.
EGC	Equal-Gain Combining.
FDMA	Frequency Division Multiple Access.
FEC	Forward Error Correction.
GA	Genetic Algorithm.
HARQ	Hybrid Automatic Repeat Request.

i.i.d	Independent Identical Distribution.
IEEE	Institute of Electrical and Electronics Engineers.
IR	Incremental Redundancy.
LDPC	Low-Density Parity-Check.
LLR	Log-Likelihood Ratio.
LoS	Line of Sight.
LTE	Long Term Evolution.
MAI	Multiple Access Interference.
MAP	Maximum A Posteriori.
MIMO	Multiple-Input Multiple-Output.
MISO	Multiple-Output Single-Input.
ML	Maximum Likelihood.
MMSE	Minimum Mean Square Error.
MRC	Maximum Ratio Combining.
MU	Multi-User.
MUD	Multi-User Detection.
NACK	Negative Acknowledgment.
PDF	Probability Density Function.
QAM	Quadrature Amplitude Modulation.
RTD	Round-Trip Delay.
SC	Selection Combining.
SDM	Space Division Multiplexing.
SDMA	Space Division Multiple Access.
SIMO	Single-Input Multiple-Output.
SINR	Signal-to-Interference-plus-Noise Ratio.
SISO	Single-Input Single-Output.
SNR	Signal-to-Noise Ratio.
SSC	Switch and Stay Combining.

SU	Single-User.
TDMA	Time Division Multiple Access.
TTO	Transmission Time Out.
V-BLAST	Vertical Bell Laboratory Layered Space-Time.
WiMAX	World Interoperability for Microwave Access.
WLAN	Wireless Local Area Network.
ZF	Zero Forcing.

Notation

$\lvert \cdot \rvert$	Absolute value.
β	Amplifying factor.
$\bar{\gamma}, \rho$	Average received SNR.
$I_0(x)$	Zero-order modified Bessel function of the first kind.
M_c	Number of bits per symbol.
$F_X(x)$	CDF of random variable X.
R_c	Code rate.
B_c	Channel coherence bandwidth.
T_c	Channel coherence time.
$(\cdot)^*$	Complex conjugate.
$\mathbb{C}$	Complex numbers.
Π^{-1}	Bit-deinterleaver.
$D_{j,j}$	(j,j)th diagonal element of matrix $\mathbf{D}$.
$\mathcal{P}_{rate}$	Dropped packet-rate.
$\mathbb{E}\{\cdot\}$	Expectation operator.
$\exp(\cdot)$	Exponential function.
$(\cdot)^{\mathrm{H}}$	Hermitian (complex conjugate transpose).
$c(\tau, t)$	Channel impulse response.
γ	Instantaneous received SNR.
Π	Bit-interleaver.

$L(c_{j,n})$	LLR for the nth coded bit of the jth transmit antenna.
$\log_2(\cdot)$	Logarithm of base 2.
$\mathbf{X}$	Matrix.
$\max(\cdot)$	Maximum operator.
$\min(\cdot)$	Minimum operator.
T_p	Packet duration.
$f(r)$	PDF of random variable r.
$P^{\text{out}}(\cdot)$	Outage probability.
$\Pr(\cdot)$	Probability.
$(\cdot)^\dagger$	Pseudo-inverse.
$\mathcal{Q}(\cdot)$	Quantization operator.
$\mathcal{R}$	Transmission rate.
ξ	Rice factor.
x	Scalar.
B_s	Signal bandwidth.
N_0	Noise power spectral density.
$\mathcal{U}_l$	Set of source nodes that are successfully decoded at relay R_l.
E_s	Average energy per symbol.
T_s	Symbol duration.
η	Throughput efficiency.
N_t	Total number of transmit antennas across all users/sources.
$(\cdot)^T$	Transpose.
$\mathbf{x}$	Vector.

Chapter 1

Introduction

Over the past two decades, wireless applications and services such as cell phones, portable computers and wireless local area networks (WLANs) have progressively become an indispensable part of our daily life. Consequently, the demands of emerging wireless services for high data rate and reliable transmission over wireless channels have also significantly escalated over time. Unfortunately, satisfying the increasing demand for high data rates is a challenging task due to several technical reasons. Firstly, the radio spectrum available for wireless services has physical and legal restrictions and therefore it is a limited resource. One needs to maximize the data rate within a given bandwidth. Secondly, increasing the transmit power of a wireless device is not an efficient approach since it will reduce its battery life and cause interference to nearby devices.

Without increasing the bandwidth or transmit power, the spectral efficiency can be greatly improved by employing multiple antennas at both ends of a link. Such systems are referred to as multiple-input multiple-output (MIMO) systems and have been broadly deployed in many wireless standards such as IEEE 802.11n WLAN [1], IEEE 802.16 world interoperability for microwave access (WiMAX) [2] and long term evolution (LTE) [3]. Transmission over the MIMO channels can be performed in different ways to achieve the spatial multiplexing gain, spatial diversity gain, or a trade-off between the spatial multiplexing gain and the spatial diversity gain [4, 5].

Reliable transmission of data packets over wireless channels suffers from multipath fading, path loss, shadowing, and interference from other transmitters. To improve the reliability, automatic repeat request (ARQ) schemes were introduced. ARQ is an error control technique widely used in wired as well as wireless communication systems to ensure reliable transmission of data packets based on feedback messages and retransmissions [6]. Error detection bits such as cyclic redundancy check (CRC) bits are added to the packet in order to determine whether the received packet contains errors. Depending on success or failure of the decoded packet, the receiver sends back a positive acknowledgment (ACK) or negative acknowledgment (NACK) message to the transmitter.

Wireless networks employing ARQ schemes can suffer from low throughput efficiency when multiple retransmissions occurs due to poor channel conditions (deep fading). In order to provide better performance and to ensure reliable communication over noisy channels, ARQ schemes can be combined with forward error correction (FEC) codes, refereed to as hybrid ARQ (HARQ). Incorporating an ARQ scheme and FEC code can result in higher system throughput efficiency than using an ARQ scheme alone [7]. HARQ schemes are mainly grouped into two categories: type-I HARQ and type-II HARQ protocols. The main difference between type-I HARQ and type-II HARQ protocols is that in type-I HARQ, when a received packet contains errors (according to the CRC), the receiver discards the erroneous packet and requests a retransmission of that packet. In type-II HARQ protocols, previous erroneously received transmissions are stored in a buffer, and are combined with the successive retransmissions to improve decoding reliability.

Type-II HARQ protocols can further be classified into two categories: Packet or Chase combining HARQ (CC-HARQ) [8] and incremental redundancy HARQ (IR-HARQ) [9]. In Chase combining, if a packet is found in error, a request is sent to the transmitter to retransmit the same packet until decoding succeeds or the number of retransmissions reaches some maximum. All versions of the same packet are com-

bined. In IR-HARQ, instead of sending the original packet in each retransmission, the transmitter sends only the additional redundancy bits. Compared with CC-HARQ, IR-HARQ is a more sophisticated HARQ protocol which requires large buffer size at the receiver and adds more complexity to the system, but it can achieve better performance [10].

In this thesis, we focus on cross-layer design between the physical layer and the data link layer of an uplink multi-user non-cooperative and cooperative system. Detection techniques are considered at the physical layer, whereas HARQ protocols are considered at the data link layer. The main objective is to design efficient signal processing techniques that can improve the error-rate performance and throughput of an uplink multi-user non-cooperative and cooperative system with reasonable computational complexity.

1.1 Research Background and Motivation

In this section, overviews of non-cooperative and cooperative communication systems employing HARQ protocols and motivation for this work are given.

1.1.1 Non-Cooperative HARQ Systems

Generally for most multiple antenna systems, e.g. vertical Bell laboratory layered space-time (V-BLAST) systems [11, 12], it is assumed that the number of receive antennas is greater than or equal to the number of transmit antennas. However, there could be times when the number of transmit antennas exceeds the number of receive antennas, leading to an overloaded (also called rank-deficient or underdetermined) system[1]. One of the key challenges of multiple antenna systems is to recover the transmitted information in the presence of co-channel interference (CCI) and noise, especially under overloaded conditions. CCI arises when multiple users transmit their information

[1]Overloaded conditions can be found in uplink multi-user communication systems, in which we have fewer receive antennas at the base station (BS) or at the access point (AP) than the number of transmitting users.

through the same channel (i.e. sharing the same space, time and frequency).

The performance of suboptimal linear detection techniques, such as zero forcing (ZF) [13] and minimum mean square error (MMSE) [13] drastically degrades in overload systems. Traditional V-BLAST algorithms [14] fail in such conditions, due to matrix singularity [15]. The optimum maximum likelihood (ML) algorithm [16] jointly detects all the transmitted signals and works well under overloaded conditions [17]. However, its computational complexity grows exponentially with the constellation size and number of transmit antennas, making it infeasible for practical applications. A generalized sphere decoding algorithm was first proposed in [18] to achieve ML performance under overloaded conditions, with a computational complexity exponential in the number of excess transmit antennas[2]. Various other suboptimal groupwise detection [15, 17] techniques and metaheuristic approaches such as ant colony optimization (ACO) [19] and genetic algorithms (GA) [20] have been proposed for overloaded systems. However, these suboptimal algorithms demand sophisticated implementation and still have quite high computational complexity. Furthermore, these algorithms often exhibit error floors at high signal-to-noise ratio (SNR) region [21].

Channel augmentation is a form of virtual MIMO techniques, where additional virtual antennas are created in order to increase the rank of the channel matrix by means of signal processing methods [22]. In [23], it was shown that the information outage rate of MIMO channels can be improved by repetitively transmitting the same spatial multiplexed signals. The received signal vectors of the same transmitted signals were stacked to create additional virtual receive antennas, thereby increasing the effective rank of the MIMO channel. Recently in [24], a linear MIMO hybrid automatic repeat request (HARQ) precoder was designed to optimize the mutual information, and joint HARQ detection was performed by stacking the received vectors from all (re-)transmissions. However, only critically loaded conditions were considered in [24]. Motivated by the improvements offered by virtual MIMO techniques, we present several

[2]The number of excess transmit antennas is equal to the difference between the number of transmit and receive antennas.

multi process HARQ schemes employing the concept of virtual receive antennas [23] for overloaded multiuser non-cooperative systems. The objective is to investigate new signal processing techniques that can improve the error-rate performance and allow us to apply suboptimal linear detectors under overloaded conditions.

1.1.2 Cooperative HARQ Relaying Systems

Advances in signal processing techniques for MIMO point-to-point links have shown tremendous improvements in capacity (multiplexing) and reliability (spatial diversity). However, due to size and cost constraints, implementation of multiple antennas may not always be feasible in small wireless devices. Recently, the concept of cooperative communication [25] has been proposed to overcome this limitation. The basic idea is to allow multiple single antenna devices to share their antennas to create a virtual multiple antenna system. Cooperation among nodes can result in a better network coverage, link reliability and improved spatial diversity gains [26, 27, 28].

The general idea behind cooperative or relay-assisted communications comes from the study of relay channels, initially proposed by van der Meulen [29, 30] and later studied extensively by Cover and El Gamal [31]. The simplest cooperative network is a three-terminal relay channel [30] consisting of a source, a relay and a destination node as illustrated in Figure 1.1. In a wireless relay channel the source node transmits its information to the destination (solid lines), which is also received by the relay node. The relay then processes the received information and forwards it to the destination (dashed lines), which combines the transmissions received by the source and relay nodes in order to decode the information. Based on the signal processing techniques adopted at the relay node, the relaying protocols are generally classified into two categories: amplify-and-forward (AF) and decode-and-forward (DF) protocols [26, 32].

In AF relaying, the relay node simply retransmits a scaled version of the information it receives from the source node. The AF or non-regenerative relaying is the simplest relaying protocol and does not require much signal processing at the relay node. The

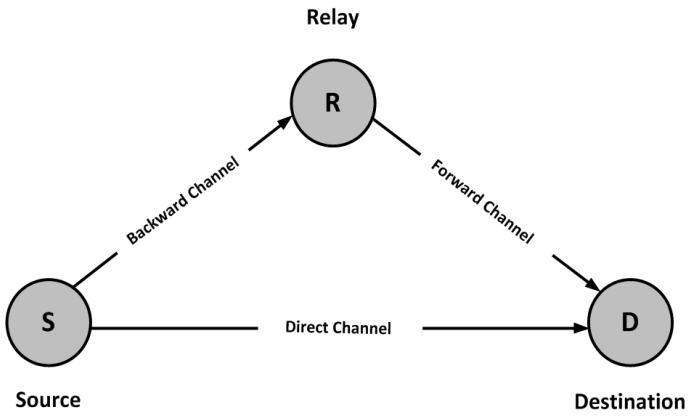


Figure 1.1: Three-terminal wireless relay channel.

main drawback of AF is that amplifying the received information will also amplify the noise at the relay node. On the other hand, in DF relaying, the relay node decodes the information it receives from the source node, re-encodes and forwards it to the destination, Hence, it is also called regenerative relaying. Depending upon the channel conditions of the source-relay link it is possible for a relay node to decode the received information in error. Thus, a retransmission of the information[3] may be required at the relay. When the source-relay channel is excellent, the DF relaying can outperform AF relaying in terms of error-rate and outage probability performance. However, the cost is increased complexity at the relay node.

Under poor channel conditions, a cooperative relaying scheme alone is not adequate to prevent packet loss and therefore, ARQ or HARQ protocols can be combined with a cooperative relaying scheme to further improve the reliability of data transmission over noisy wireless channels [33]. Integrating these two techniques was originally proposed in [34]. It was shown that cooperative HARQ relay networks provide many advantages in terms of delay, throughput and energy consumption compared to conventional point-to-point HARQ networks. Since then, interest in cooperative HARQ relaying systems has rapidly grown among researchers. Numerous retransmission schemes have been developed for a single-relay network [33, 35, 36, 37] and for multi-relay networks [38,

[3]Depending on the type of retransmission protocol considered, the retransmitted information can be a copy of the original information or new information containing only parity bits.

39, 40, 41].

In [33] a combined hybrid strategy called H^2-ARQ-relaying was proposed in which both the retransmission protocols and the relaying schemes are hybrid. In [35], an analytical methodology for the evaluation of outage probability of a DF ARQ was proposed in which the relay forwards Alamouti [42] based retransmission signals. Harsini *et al.* [36] analyzed the throughput and packet loss-rate performance of non-cooperative and cooperative type-II HARQ protocols in adaptive rate wireless network with time-correlated fading channels. In [37], the authors analyzed a cooperative HARQ transmission with incremental redundancy over a block fading channel. Dynamic programming was used to find the optimal rate adaptation policy to optimize the throughput of cooperative HARQ transmission. A distributed truncated ARQ protocol for an AF cooperative system was proposed in [38], where a retransmission was made by the source or the best instantaneous relay node. A delay analysis of a cooperative truncated HARQ system with opportunistic relaying for both AF and DF was presented in [39]. In [40], the throughput for cooperative ARQ and type-I HARQ protocols with opportunistic relaying was evaluated. Maham *et al.* [41] analyzed the outage probability of the multi-relay HARQ system with opportunistic relaying scheme in DF mode. All these studies assume that there is always a direct link between source and destination.

Unfortunately, due to large physical separation or path loss, a direct link may not always be available. In the absence of a direct link, relay(s) adopting a half-duplex (meaning they can either transmit or receive at a given time) DF relaying scheme with HARQ protocol will request retransmission whenever they receive an erroneous packet. Multi-hop HARQ DF relay networks with no direct link were studied in [43, 44, 45, 46, 47]. The diversity-multiplexing-delay tradeoffs for various multi-hop ARQ protocols were analyzed in [43]. In [44], the authors proposed a method to handle the maximum number of transmissions between the broadcast link and relay link by taking into account the end-to-end error probability and spectral efficiency. Zheng et al. [45] presented an analytical framework for a cooperative DF relaying system with general

hop-by-hop HARQ transmission. The authors derived an upper bound on the block-error-rate assuming that the transmission from the source to the destination is hop-by-hop through the relay. In [46], the optimization of a two-hop DF relaying system with HARQ was studied in order to maximize the long-term average transmission rate. In [47], the authors analyzed the energy efficiency of a wireless multiple relays network with IR-HARQ protocol and distributed cooperative beamforming. An important limitation of these and similar studies is that they have considered only a single source that is assisted by one or multiple relays. The extent to which the above results may be generalized to multiple source scenarios remains unclear.

This motivates our work in this thesis, where we focus on multi-source (re-)transmission techniques for a two-hop multi-relay DF relaying system. To the best of our knowledge this is the first study looking at HARQ for a multi-source scenario in the absence of a direct link.

1.2 Thesis Contributions and Outline

The main contributions of this thesis are summarized as follows:

- A new multi-user transmission approach that combines HARQ and a virtual receive antenna concept is proposed for overloaded multi-user non-cooperative systems. This basic idea is to treat the HARQ retransmission information as if it was arriving to an additional virtual receive antennas. Based on the proposed idea, two detection schemes are presented. In Scheme-I, users with correctly decoded packets remain idle, whereas users with erroneous packets retransmit their packets in subsequent frames. In Scheme-II, users with correctly decoded packets may transmit new information instead of remaining idle. These schemes allow us to apply suboptimal linear MUD algorithms rather than employing high computational complexity detection algorithms to handle overloaded conditions.

- A multi process HARQ retransmission scheme is proposed for multi-source multi-

relay DF relaying networks with no direct link, a topic that has not been investigated in the literature. The proposed scheme that utilizes virtual antennas to achieve increased receive diversity is allowed simultaneous retransmissions from source node(s) of packets detected in error. This results in improved throughput compared to traditional orthogonal (time division) retransmissions since fewer time slots are required by the proposed scheme to complete the HARQ round. Moreover, to avoid unnecessary retransmissions upon completion of each (re-)transmission round from the source nodes, relays are allowed to exchange their decoding outcomes.

- A novel forwarding strategy with minimum overhead is proposed for a two-hop DF relaying system. The basic idea behind the proposed strategy is to assign each relay one priority source node. The relay nodes are allowed to simultaneously transmit the data packets of source nodes to the destination, thereby improving the throughput. Based on the exchange of decoding outcomes (ACK/NACK) among relay nodes, the relays can independently decide to swap their priority source nodes. Switching of priority source nodes at the relays does not require any centralization or participation either from the sources or the destination.

- The performance of the two-hop DF relaying system with HARQ is also analyzed. First, we derive the expressions of outage probabilities of each transmission phase and then the end-to-end (e2e) outage probabilities for both the proposed and orthogonal time division (re-)transmissions schemes. The e2e outage probability is an important parameter to determine the reliability of a multi-hop DF relaying system in the absence of a direct link. The renewal-reward theorem [48] is used to derive throughput expressions for the proposed schemes. For comparison purpose, the throughput performance of orthogonal time division (re-)transmissions is also evaluated. Finally, the analytical results are compared and confirmed with Monte-Carlo simulation results. It is shown that the proposed schemes achieve higher throughput compared to the traditional (time division) (re-)transmissions at the

cost of a slight performance degradation in outage probability due to CCI.

Papers submissions and publications:

- Z. Rauf, P. A. Martin and D. P. Taylor, "Multiuser detection of overloaded systems employing HARQ," in Proc. IEEE Int'l. Conf. on Commun. Systems (ICCS), pp. 300-304, 21-23 Nov. 2012.
- Z. Rauf, P. A. Martin and D. P. Taylor, "Multi-source multi-relay cooperative systems employing HARQ," IEEE Trans. Veh. Tech., under revision , 2013.

The remainder of this thesis is organized as follows:

Chapter 2 presents some basic concepts and background material related to this thesis. It starts with a brief overview of wireless channel characteristics and different statistical models. Then, the gains achieved by using multiple antennas at the transmitter and/or at the receiver are discussed. After a brief description of point-to-point and multi-user MIMO system models, various multi-user detection techniques are summarized. This is then followed by a brief overview of different types of ARQ and HARQ protocols used to ensure reliable transmission. At the end of this chapter, several cooperative network configurations and relaying schemes are presented.

Chapter 3 deals with the multi-user detection and HARQ techniques and their performance under overloaded conditions. First, the system model of an uplink multi-user non-cooperative system with HARQ retransmissions is presented. Then, the two proposed retransmission schemes that combine HARQ with the use of virtual receive antennas [23] are described. This combination results in a simple and novel multi-user transmission approach which works well under overloaded conditions. Finally, simulation results are presented, showing the performance gains achieved by the proposed schemes in terms of BER and dropped packet rate.

Chapter 4 focuses on multi process (re-)transmission schemes for a multi-source and multi-relay cooperative network. Firstly, the system model for a two-hop DF relaying system is described. Then, the proposed retransmission and forwarding schemes during

the broadcast and relay phase, respectively, are discussed. Exchange of decoding among relay nodes is allowed which results in an improved retransmission rate as demonstrated by the simulation results.

Chapter 5 develops the outage probability and throughput performance analysis of the proposed two-hop DF relaying system discussed in Chapter 4.

Chapter 6 summarizes the whole thesis and outlines several possible avenues for future research.

Chapter 2

Background

The objective of this chapter is to provide important background knowledge required for the subsequent chapters. In Section 2.1 basic properties of multipath wireless channels and their modeling are described. Section 2.2 presents a brief overview of multiple antenna wireless systems, gains achieved by employing multiple antennas at the transmitter and/or at the receiver, and system models for single and multi-user multiple-input and multiple-output (MIMO) communication links. Various optimal and suboptimal multi-user detection (MUD) techniques for uncoded and coded transmissions are discussed in Section 2.3. Section 2.4 briefly outlines different retransmission techniques. Section 2.5 introduces cooperative relaying systems and various strategies adopted at the relay node(s).

2.1 Wireless Fading Channels

Compared to wired communication, reliable transmission over wireless channels is challenging as it suffers from many channel impairments. The fading effects that characterize wireless communication over such channels can be classified into two different categories, namely large-scale fading (includes path loss and shadowing) and small-scale fading also referred to as multipath fading [49]. The path loss is the result of signal power attenuation resulting from propagation over large distances, and shadowing is caused by signal power loss due to large objects in the environment, such as

buildings, trees and hills. Small-scale fading or multipath fading is due to constructive and destructive combination of multiple attenuated versions of the transmitted signal arriving at the receiver from a number of different paths, which experience differences in attenuation, delay and phase shift.

Large-scale fading occurs over large transmitter-receiver separation distances and is more related to the issues such as cell-site planning [49]. On the other hand, small-scale fading occurs over short distances and is more relevant to the design aspects of reliable and efficient communication systems [49], which are also the focus of this thesis.

2.1.1 Statistical Models for Fading Channels

Statistical models for the small-scale fading channels are presented in this section.

Rayleigh Fading Channel

A Rayleigh fading channel is commonly used to model highly built-up environments, where there is no line of sight (LoS) path between the transmitter and receiver, and there are many objects on the path that scatter the transmitted signal. When there is no dominant propagation along a LoS path and the transmitted signal is received at the receiver via different paths, the channel impulse response $c(\tau, t)$ at a delay τ and time instant t can be modeled as a zero-mean complex-valued Gaussian process [50]. In this case the envelope $|c(\tau, t)|$ has a Rayleigh distribution and the channel is said to be a Rayleigh fading channel. The probability density function (PDF) of the Rayleigh distribution is expressed as [51]

$$f(r) = \begin{cases} \frac{r}{\sigma^2} \exp\left(-\frac{r^2}{2\sigma^2}\right) & \text{for} \quad r \geq 0 \\ 0 & \text{otherwise,} \end{cases} \tag{2.1}$$

where r is the envelope of the received signal and σ^2 is the mean received power.

Rician Fading Channel

In the presence of a LoS path between the transmitter and the receiver, the channel impulse response can be modeled as a non zero-mean complex-valued Gaussian process [50]. In this case the received signal can be statistically described by the Rician distribution and the channel is said to be a Rician fading channel. The PDF of the Rician distribution is given by [51]

$$f(r) = \begin{cases} \frac{r}{\sigma^2} \exp\left(-\frac{\left(r^2+A^2\right)}{2\sigma^2}\right) I_0\left(\frac{rA}{\sigma^2}\right) & \text{for} \quad r \geq 0, A \geq 0 \\ 0 & \text{otherwise,} \end{cases} \tag{2.2}$$

where A is the peak magnitude of the LoS signal component and $I_0(x)$ is the modified zero-order Bessel function of the first kind, defined as [52]

$$I_0(x) = \frac{1}{2\pi} \int_0^{2\pi} \exp\big(x\cos(\theta)\big) d\theta. \tag{2.3}$$

The Rician distribution is usually characterized by the Rice factor ξ, which is defined as the ratio between the LoS signal power and the non-LoS multipath signal power. It is given by

$$\xi = \frac{A^2}{2\sigma^2}. \tag{2.4}$$

As we can see, when $\xi = 0$, the Rician distribution (2.2) reduces to a Rayleigh distribution (2.1).

2.1.2 Fast and Slow Fading

The classification of the fading channel as fast or slow depends on the relationship between the symbol duration T_s (seconds) of the transmitted signal and the coherence time T_c (seconds) of the channel. The coherence time determines how fast the channel is changing in time.

- In fast fading, the coherence time of the channel is smaller than the symbol

period of the transmitted signal. Fast fading (also called time-selective fading) is the result of rapid changing of the channel impulse response during the span of a symbol.

- In slow fading, the coherence time of the channel is greater than the symbol period of the transmitted signal, and thus the channel can be assumed to be static over multiple symbol durations. The channel is said to be quasi-static, if the channel remains constant during one packet[1] transmission. In this case, $T_c = T_p$, where T_p denotes the time required to transmit a packet. If we have F packets to transmit and the channel remains constant during F transmissions, i.e. $T_c = FT_p$, the channel is said to be block fading.

2.1.3 Frequency-Flat and Frequency-Selective Fading

Depending on the relation between the bandwidth B_s (Hz) of the transmitted signal and the coherence bandwidth B_c (Hz) of the channel, the fading channel can be classified as frequency-flat or frequency-selective fading.

- In frequency-flat fading, the coherence bandwidth of the channel is greater than the bandwidth of the transmitted signal. Hence, all frequency components of the transmitted signal will experience the same attenuation. Frequency-flat fading channels are also referred to as narrowband channels, since the bandwidth of the transmitted signal is narrow compared to the coherence bandwidth of the channel [53].

- In frequency-selective fading, the bandwidth of the channel is smaller than the bandwidth of the transmitted signal. Therefore, different frequency components of the transmitted signal will experience different gains and phase shifts. Frequency-selective fading channels are also known as wideband channels, since the bandwidth of the transmitted signal is wide compared to the coherence bandwidth of the channel [53].

[1] A packet or a frame usually consists of more than one symbol.

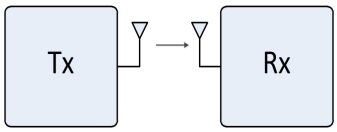
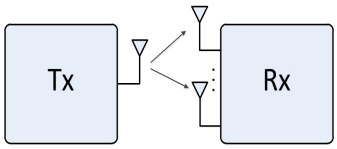
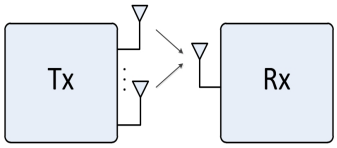
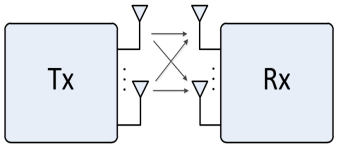
Multiple Antenna Wireless Systems		
SISO	single-input single-output systems have only one antenna at the transmitter and receiver.	
SIMO	single-input multiple-output systems have multiple antennas at the receiver and one antenna at the transmitter.	
MISO	multiple-input single-output systems have multiple antennas at the transmitter and one antenna at the receiver.	
MIMO	multiple-input multiple-output systems have multiple antennas at the transmitter and receiver.	

Table 2.1: Antenna configuration of a transmitter (Tx) and receiver (Rx) in a wireless system.

2.2 Overview of Multiple Antenna Communication Systems

Wireless systems consisting of a transmitter, a radio channel and a receiver are generally characterized by the number of inputs and outputs as illustrated in Table 2.1. Single-input single-output (SISO) is the simplest configuration with a single antenna at both ends of a link. When multiple antennas are used only at the receiver or only at the transmitter, the wireless systems are denoted as single-input multiple-output (SIMO) systems or multiple-output single-input (MISO) systems, respectively. Employing multiple antennas at both ends of the point-to-point communication link re-

sults in multiple-input multiple-output (MIMO) systems. The main gains[2] achieved by employing multiple antennas include [55]:

- **Array Gain:** Signals from multiple antennas at the receiver can be combined coherently to obtain gain in the average Signal-to-Noise Ratio. Such gain is usually referred to as array gain. The average SNR grows proportionally to the number of receive antennas [55]. Combining methods at the receiver can generally be classified into four types: selection combining (SC), switch and stay combining (SSC), maximum ratio combining (MRC) and equal-gain combining (EGC). A detailed description of combining methods can be found in [56]. Array gain can also be obtained using multiple antennas at the transmitter[3].

- **Diversity Gain:** Diversity is an effective approach to combat the effects of multipath fading. Diversity techniques can be classified into three different domains: time, frequency and spatial diversity. Time diversity is most applicable to the cases where the symbol period of the transmitted signal is greater than the coherence time of the channel (i.e. time-selective fading) [57]. Frequency diversity is effective to the frequency-selective fading channels, where the symbol bandwidth of the transmitted signal is greater than the coherence bandwidth of the channel [57]. Spatial diversity (also known as antenna diversity) is widely utilized in wireless communication systems and can be categorized into receive diversity and transmit diversity, depending on whether it is applied at the receiver or at the transmitter side.

  - Receive Diversity: In receive diversity, the receiver combines the independently faded versions of the same transmitted signal from different receive antennas, so that the combined signal suffers less attenuation (fading) than the received signal at any one antenna [58]. In this technique, the number of

[2] In general, exploiting these gains comes at the price of drastically increased signal processing complexity, especially at the receiver side. However, achieving all gains at once may not always be feasible [54].

[3] Note that the array gain exploitation requires perfect knowledge of the channel state information (CSI) at the transmitter or receiver or both.

independent diversity branches (also called diversity order) is equal to the number of receive antennas.

- Transmit Diversity: In transmit diversity, space-time coding schemes, such as the Alamouti scheme [42] can be employed to achieve transmit diversity without knowledge of the channel at the transmitter. For transmit diversity, the diversity order is equal to the number of transmit antennas.

- **Multiplexing Gain:** Spatial diversity gain can be acquired by employing multiple antennas either at the transmitter or at the receiver side. On the contrary, spatial multiplexing gain requires multiple antennas at both ends of a link. An increase in data rate without any additional power expenditures can be obtained by multiplexing the transmitted data streams among different antennas [55]. This increase in data rate is proportional to the minimum of the number of transmit antennas and the number of receive antennas.

  In spatial multiplexing, several data streams are simultaneously transmitted from multiple antennas over the MIMO channel [55]. The receiver employs some kind of multi-signal detection algorithm to recover these individual data streams.

A point-to-point MIMO system also known as single-user MIMO (SU-MIMO) system can be extended to a multi-user MIMO (MU-MIMO) system. A cellular system is an example of such a MIMO system, where multiple users each with one or more antennas communicate with a base station (BS) with multiple antennas [59]. For the uplink or multiple access channels, the BS is the receiver and users are the transmitters, while for the downlink or broadcast channels, the roles are reversed.

Simultaneous transmission of several data streams on different transmit antennas over the point-to-point communication link is known as space division multiplexing (SDM). For the multi-user scenarios, different users simultaneously transmit at the same time and frequency. Such a technique is referred to as space division multiple access (SDMA) [59]. This thesis concentrates on MU-MIMO systems, although point-to-point systems are discussed in general. In particular, we focus on uplink channels

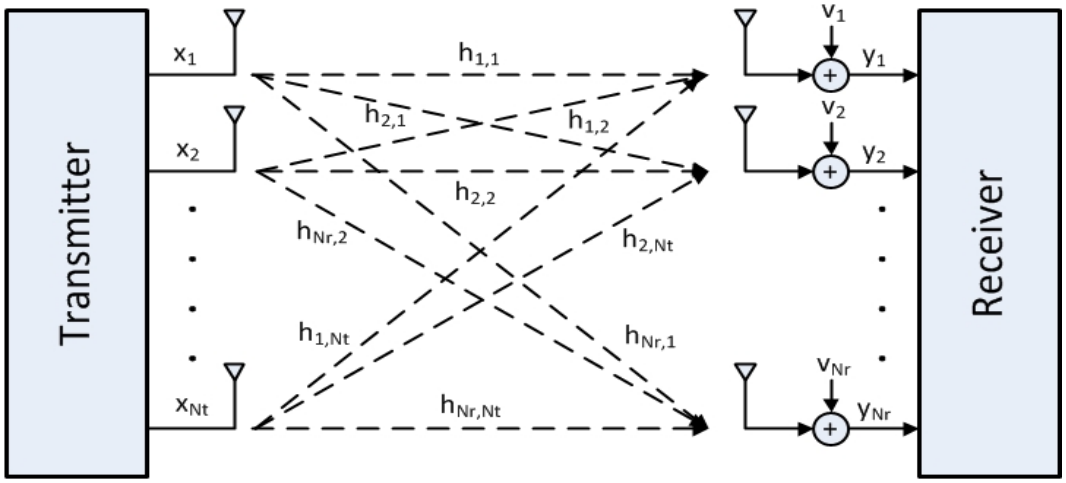


Figure 2.1: Wireless point-to-point MIMO link, where the transmitter is equipped with N_t antennas and the receiver is equipped with N_r antennas.

where each user is equipped with one antenna. Section 2.2.1 and Section 2.2.2 present the system models for the point-to-point MIMO system and the uplink MU-MIMO system, respectively.

2.2.1 Point-to-Point MIMO System Model

Let us consider a single-user MIMO system with N_t transmit antennas and N_r receive antennas, as shown in Figure 2.1. Throughout this thesis, the channel is assumed to exhibit quasi-static, frequency-flat Rayleigh fading, meaning the channel coefficients are constant over a packet of length L and vary independently between packets. Let $h_{i,j}$ denote the channel coefficient from transmit antenna j to receive antenna i. These are modeled as independent identically distributed (i.i.d) complex Gaussian random variables with zero-mean and unit variance. At symbol interval t, the complex symbol vector $\mathbf{x}(t) = \big(x_1(t), \cdots, x_j(t), \cdots, x_{N_t}(t)\big)^T$ drawn from an alphabet $\mathcal{A} = (a_1, a_2, \cdots, a_C)$, is transmitted over the MIMO channel from the N_t transmit antennas. Here $(\cdot)^T$ denotes the transpose and C represents the alphabet size. It is given by

$$C = |\mathcal{A}| = 2^{M_c}, \tag{2.5}$$

where M_c is the number of bits per symbol. The complex received signal at antenna i can be expressed as

$$y_i(t) = \sum_{j=1}^{N_t} h_{i,j}(t)x_j(t) + v_i(t), \qquad i = 1, 2, \ldots, N_r, \quad j = 1, 2, \ldots, N_t, \tag{2.6}$$

where the term $v_i(t)$ represents the additive noise. The scalar relationship given in (2.6) can be expressed in vector-matrix form as

$$\mathbf{y}(t) = \mathbf{H}(t)\mathbf{x}(t) + \mathbf{v}(t), \qquad t = 1, 2, \ldots, L, \tag{2.7}$$

where $\mathbf{y}(t) \in \mathbb{C}^{N_r \times 1}$ is the complex received signal vector, $\mathbf{H}(t) \in \mathbb{C}^{N_r \times N_t}$ represents the complex channel matrix given by

$$\mathbf{H}(t) = \begin{bmatrix} h_{1,1}(t) & h_{1,2}(t) & \cdots & h_{1,N_t}(t) \\ h_{2,1}(t) & h_{2,2}(t) & \cdots & h_{2,N_t}(t) \\ \vdots & \vdots & \ddots & \vdots \\ h_{N_r,1}(t) & h_{N_r,2}(t) & \cdots & h_{N_r,N_t}(t) \end{bmatrix}, \tag{2.8}$$

and $\mathbf{v}(t) \in \mathbb{C}^{N_r \times 1}$ is a complex additive white Gaussian noise (AWGN) vector whose components each have zero-mean and variance σ_v^2. At the receiver, a hard-output detection algorithm is employed to compute the estimates of the transmitted symbol vector $\mathbf{x}(t)$ that is corrupted by the wireless fading channel and the AWGN noise.

For coded transmission over point-to-point MIMO links, a packet $\mathbf{b}$ consisting of K bits (including the information bits and cyclic redundancy check (CRC) bits) is first encoded by a channel encoder of rate $R_c = K/N$, resulting in a codeword $\mathbf{c}$ of length N coded bits. The coded bits are then re-ordered by a random bit-interleaver (Π). For a given symbol interval, the symbol vector $\mathbf{x}(t)$ is obtained by mapping $\mathbf{c} = (c_1, \cdots, c_n, \cdots, c_{N_t M_c})$ interleaved coded bits . The symbol vector $\mathbf{x}(t)$ is then transmitted over the MIMO channel from the N_t transmit antennas.

The block diagram of a coded point-to-point MIMO system is depicted in Figure

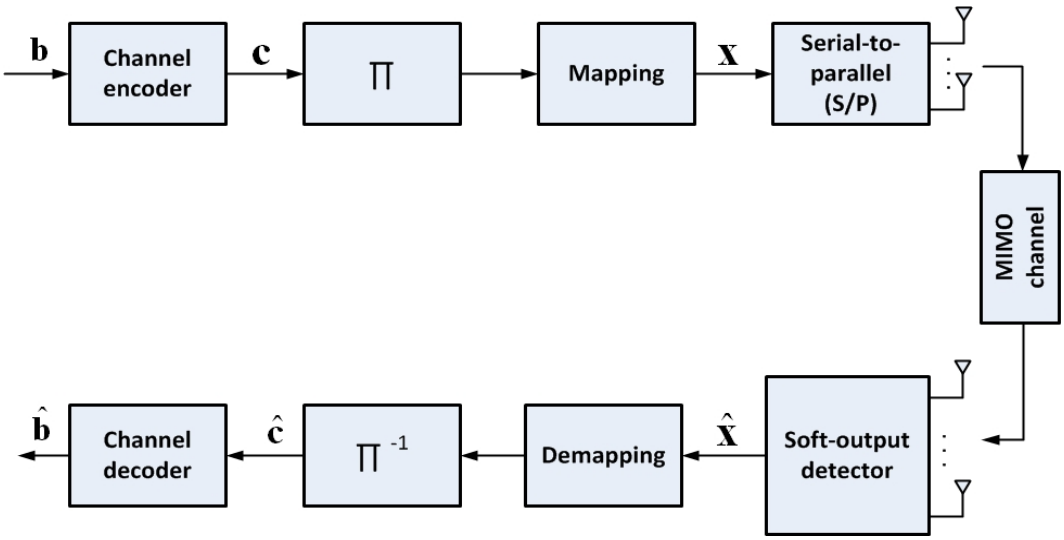


Figure 2.2: Block diagram of a coded point-to-point MIMO system.

2.2. At the receiver, a soft-output detector jointly detects the coded bits contained in the received signal vectors $\mathbf{y}(t)$. The soft estimates are passed to the bit-deinterleaver (Π^{-1}) and then to the channel decoder. The channel decoder computes hard estimates of the information bits $\hat{\mathbf{b}}$.

2.2.2 Multi-User MIMO System Model

Figure 2.3 shows the block diagram of an uplink multi-user communication system with U co-channel users each communicating with a centralized receiver (i.e. a BS or an access point) having N_r receive antennas. Let N_t denotes the total number of transmit antennas across all users. Since each user is equipped with a single antenna (assumed), therefore, we let $N_t = U$. The input-output relation for the uplink multi-user channels is similarly defined as a point-to-point MIMO link except that each input corresponds to a user. For simplicity, it is assumed here that all co-channel users are perfectly synchronized and transmit simultaneously. For a given symbol interval, the complex symbol $x_j(t)$ of the jth user drawn from an alphabet $\mathcal{A}$ is transmitted over the frequency-flat Rayleigh fading channel from the jth antenna. The complex received

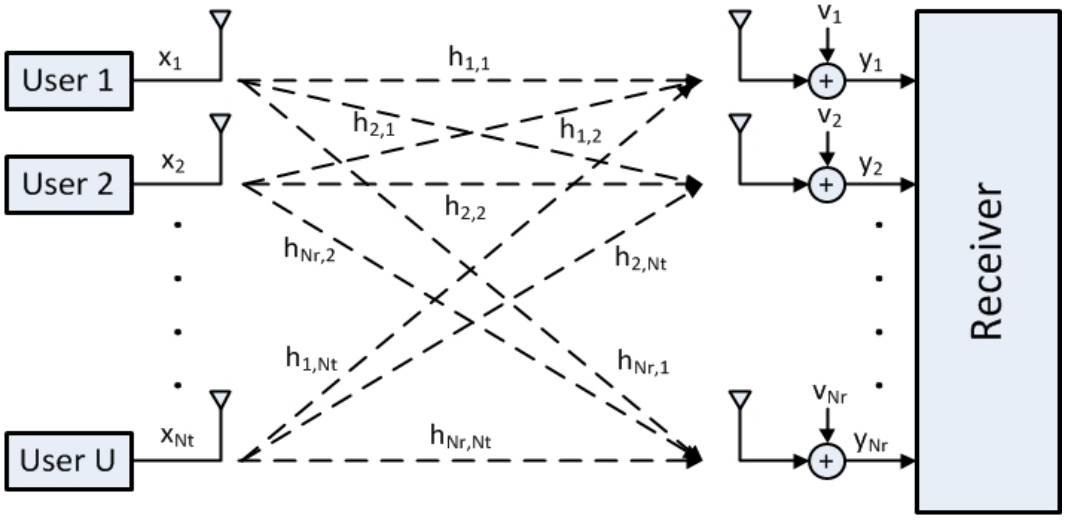


Figure 2.3: Block diagram of an uplink multi-user MIMO system, where each user is equipped with a single antenna and the centralized receiver is equipped with N_r antennas.

signal vector $\mathbf{y}(t) \in \mathbb{C}^{N_r \times 1}$ is written as

$$\begin{aligned} \mathbf{y}(t) &= \sum_{j=1}^{N_t} \mathbf{h}_j(t) x_j(t) + \mathbf{v}(t) \\ &= \mathbf{H}\mathbf{x}(t) + \mathbf{v}(t), \qquad t = 1, 2, \ldots, L, \end{aligned} \tag{2.9}$$

where $\mathbf{h}_j(t)$ is the jth column of the channel matrix $\mathbf{H}(t)$ of size $N_r \times N_t$ whose entry, $h_{i,j}(t)$, corresponds to the channel coefficient between the ith receive antenna and the jth user and $\mathbf{x}(t) = \big(x_1(t), \cdots, x_j(t), \cdots, x_{N_t}(t)\big)^T$ is the overall transmitted symbol vector.

For coded transmission over multiple access channels, the information bits of each user are separately encoded by the channel encoders. Symbol $x_j(t)$ containing M_c coded bits is obtained by mapping $\mathbf{c}_j = (c_{j,1}, \cdots, c_{j,n}, \cdots, c_{j,M_c})$ interleaved coded bits. At the receiver, the soft estimates produced by the detector are passed to bit-deinterleavers and then to the channel decoders of each user.

For simplified notation and convenience, the dependency on index t is dropped for

the remainder of the thesis. Thus, equations (2.7) and (2.9) can be re-written as [60]

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{v}. \tag{2.10}$$

The above equation represents the input-output relation of the MIMO transmission over a single symbol interval. When the channel is constant over a packet of length L, the input-output relation of the MIMO channel can be expressed as [60]

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{V}, \tag{2.11}$$

where $\mathbf{Y} \in \mathbb{C}^{N_r \times L}$, $\mathbf{X} \in \mathbb{C}^{N_t \times L}$ and $\mathbf{V} \in \mathbb{C}^{N_r \times L}$.

2.3 Multi-User Detection Techniques

Multi-user detection deals with the joint detection of data signals that are transmitted from different users over a common channel. MUD was first studied in the context of code division multiple access (CDMA) by Verdú [61]. More recently, MUD techniques have also been applied to the joint detection of several signals transmitted over a MIMO channel. For a spatial multiplexing point-to-point MIMO system, the term multiple users corresponds to the multiple information streams transmitted from different antennas of a single user [62].

In the following sections, the optimal and various suboptimal MUD algorithms are briefly discussed for uncoded and coded transmissions over multiple access channels (Section 2.2.2). The implementation of these MUD algorithms for point-to-point MIMO systems is conceptually straightforward.

2.3.1 Hard-Output Detection

Detectors that compute the hard estimates of the transmitted symbol vector are referred to as hard-output detectors [54].

Maximum A Posteriori (MAP) and Maximum Likelihood (ML) Detection

The optimal MUD algorithm maximizes the probability of a correct estimate, $\Pr\big(\hat{\mathbf{x}} = \mathbf{x}\big)$, by performing an exhaustive search over all possible transmitted symbols. Let $\Pr\big(\hat{\mathbf{x}} = \mathbf{x}|\mathbf{y}, \mathbf{H}\big)$ denote the conditional probability that it has correctly estimated $\mathbf{x}$ given $\mathbf{y}$ and $\mathbf{H}$. Then, the MAP detection that maximizes $\Pr\big(\hat{\mathbf{x}} = \mathbf{x}\big)$ can be expressed as

$$\hat{\mathbf{x}}_{\text{MAP}} = \arg \max_{\mathbf{x} \in \mathcal{A}^{N_t}} \Pr\big(\hat{\mathbf{x}} = \mathbf{x}|\mathbf{y}, \mathbf{H}\big), \tag{2.12}$$

where $\hat{\mathbf{x}}_{\text{MAP}}$ is the MAP estimate for the transmitted symbol vector $\mathbf{x}$. When applying the Bayes's rule to (2.12), the MAP detection can be written as

$$\hat{\mathbf{x}}_{\text{MAP}} = \arg \max_{\mathbf{x} \in \mathcal{A}^{N_t}} \left\{ \frac{f\big(\mathbf{y}|\hat{\mathbf{x}} = \mathbf{x}, \mathbf{H}\big) \Pr\big(\hat{\mathbf{x}} = \mathbf{x}\big)}{f(\mathbf{y})} \right\}, \tag{2.13}$$

where $f\big(\mathbf{y}|\hat{\mathbf{x}} = \mathbf{x}, \mathbf{H}\big)$ is the conditional PDF of $\mathbf{y}$ given $\mathbf{x}$ and $\mathbf{H}$ and $\Pr\big(\hat{\mathbf{x}} = \mathbf{x}\big)$ is the *a priori* probability. The term $f(\mathbf{y})$ does not depend on $\mathbf{x}$, hence it can be omitted, resulting in

$$\hat{\mathbf{x}}_{\text{MAP}} = \arg \max_{\mathbf{x} \in \mathcal{A}^{N_t}} \Big\{ f\big(\mathbf{y}|\hat{\mathbf{x}} = \mathbf{x}, \mathbf{H}\big) \Pr\big(\hat{\mathbf{x}} = \mathbf{x}\big) \Big\}. \tag{2.14}$$

For the system model in (2.10) with AWGN, $\mathbf{v}$, the conditional PDF $f\big(\mathbf{y}|\hat{\mathbf{x}} = \mathbf{x}, \mathbf{H}\big)$ is given by [63]

$$f\big(\mathbf{y}|\hat{\mathbf{x}} = \mathbf{x}, \mathbf{H}\big) = \frac{1}{(\pi \sigma_v^2)^{N_r}} e^{-\frac{1}{\sigma_v^2} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2}. \tag{2.15}$$

This leads to the MAP detection rule

$$\hat{\mathbf{x}}_{\text{MAP}} = \arg \min_{\mathbf{x} \in \mathcal{A}^{N_t}} \Big\{ \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 - \log \Pr\big(\hat{\mathbf{x}} = \mathbf{x}\big) \Big\}. \tag{2.16}$$

If all the transmitted symbols are equiprobable, then the MAP detection rule reduces to the ML detection rule

$$\hat{\mathbf{x}}_{\text{ML}} = \arg \min_{\mathbf{x} \in \mathcal{A}^{N_t}} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2. \tag{2.17}$$

The complexity of MAP and ML detection grows exponentially with the number of transmit antennas and the alphabet size $|\mathcal{A}|$. Hence, many suboptimal detection algorithms have been proposed in the literature in order to avoid such computational complexity.

Linear Detection

In this section, two well known sub optimum linear detectors, namely, zero forcing (ZF) and minimum mean square error (MMSE) detection are presented. Both detectors have a polynomial computational complexity with respect to the number of transmit antennas [64]. However, the cost is performance loss compared to MAP or ML performance.

Zero Forcing Detection

The ZF detector suppresses the interference caused by the co-channel signals by multiplying the pseudo-inverse of the channel matrix $\mathbf{H}^{\dagger} = (\mathbf{H}^{\mathrm{H}}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{H}}$ with the channel output $\mathbf{y}$, such that

$$\begin{aligned} \tilde{\mathbf{x}}_{\mathrm{ZF}} &= (\mathbf{H}^{\mathrm{H}}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{H}}\mathbf{y} \\ &= \mathbf{x} + \mathbf{H}^{\dagger}\mathbf{v}. \end{aligned} \tag{2.18}$$

The superscripts $(\cdot)^{\dagger}$ and $(\cdot)^{\mathrm{H}}$ denote the pseudo-inverse and Hermitian transpose, respectively. The hard estimate for the jth transmit antenna[4] is obtained by the quantization operation [65]:

$$\hat{x}_{j,\mathrm{ZF}} = \mathcal{Q}(\tilde{x}_j), \qquad \text{for} \quad j = 1, 2, \ldots N_t, \tag{2.19}$$

where $\tilde{x}_j = (\tilde{\mathbf{x}}_{\mathrm{ZF}})_j$ is the soft ZF estimate of the jth transmit antenna at the output of the ZF filter and $\mathcal{Q}(\cdot)$ is the quantization operator. The main drawback of a ZF detector is that it enhances the background noise when $\mathbf{H}$ is near singular [13], resulting

[4]Since each user is equipped with a single antenna, each transmit antenna corresponds to a single user.

in performance degradation.

Minimum Mean Square Error Detection

The MMSE detector minimizes the mean square error (MSE) between the transmitted sequence $\mathbf{x}$ and the output of the linear filter. The MSE is given by [13]

$$\text{MSE} = E\left[\|\mathbf{x} - \mathbf{W}^{\text{H}}\mathbf{y}\|^2\right], \tag{2.20}$$

where $\mathbf{W} = (\mathbf{H}^{\text{H}}\mathbf{H} + \sigma_v^2\mathbf{I})^{-1}\mathbf{H}^{\text{H}}$ is a $N_r \times N_t$ linear filter and $\mathbf{I}$ is the identity matrix. Thus, the linear estimate of the transmitted sequence $\mathbf{x}$ is

$$\begin{aligned} \tilde{\mathbf{x}}_{\text{MMSE}} &= \mathbf{W}\mathbf{y} \\ &= (\mathbf{H}^{\text{H}}\mathbf{H} + \sigma_v^2\mathbf{I})^{-1}\mathbf{H}^{\text{H}}\mathbf{y}. \end{aligned} \tag{2.21}$$

Finally, the hard decision $\hat{x}_{j,\text{MMSE}}$, $j = 1, 2, \ldots N_t$, is obtained by using the quantization operator $\mathcal{Q}$. The MMSE detector takes into account both the background noise and the CCI. This yields improved performance compared to the ZF approach. As the SNR tends to infinity, the MMSE and ZF performances converge [13].

2.3.2 Soft-Output Detection

The soft-output detector calculates the log-likelihood ratio (LLR) of the transmitted bit stream. The resulting soft estimates are fed to the channel decoder of each user for iterative decoding.

Optimum Soft ML Detector

A soft-output ML detector calculates the LLR for all coded bits from the received signal vectors. The LLR for the nth coded bit of the jth transmit antenna, $c_{j,n}$, with

$n = 1, 2, \ldots, N$, is given by [63]

$$\begin{aligned} L(c_{j,n}) &= \log\left(\frac{\Pr\left(c_{j,n}=1|\mathbf{y},\mathbf{H}\right)}{\Pr\left(c_{j,n}=0|\mathbf{y},\mathbf{H}\right)}\right), \qquad j = 1, 2, \ldots, N_t, \\ &= \log\left(\frac{\sum\limits_{\mathbf{x}\in\chi_n^1} exp\left(-\frac{1}{\sigma_v^2}\|\mathbf{y}-\mathbf{H}\mathbf{x}\|^2\right)}{\sum\limits_{\mathbf{x}\in\chi_n^0} exp\left(-\frac{1}{\sigma_v^2}\|\mathbf{y}-\mathbf{H}\mathbf{x}\|^2\right)}\right), \end{aligned} \tag{2.22}$$

where χ_n^1 and χ_n^0 denote the sets of transmit symbol vectors whose nth bit equals 1 and 0, respectively (note that $\mathcal{A}^{N_t} = \chi_n^1 \cup \chi_n^0$ [63]). Applying the max-log approximation to (2.22) results in

$$L(c_{j,n}) = \frac{1}{\sigma_v^2}\left(\min_{\mathbf{x}\in\chi_n^0} \|\mathbf{y}-\mathbf{H}\mathbf{x}\|^2 - \min_{\mathbf{x}\in\chi_n^1} \|\mathbf{y}-\mathbf{H}\mathbf{x}\|^2\right). \tag{2.23}$$

The complexity of (2.23) for LLR computation is exponential in the number of transmit antennas and the constellation size.

Suboptimal Linear Soft Detectors

Now we review ZF and MMSE soft output suboptimal detectors [63] to calculate the LLRs for each transmit antenna separately. First we look at the ZF-based detector. The received vector after the ZF filter is given by

$$\tilde{\mathbf{x}}_{ZF} = \mathbf{W}\mathbf{y} = \mathbf{x} + \tilde{\mathbf{v}}, \tag{2.24}$$

where $\mathbf{W} = \mathbf{H}^{\dagger} = \left(\mathbf{H}^{\mathrm{H}}\mathbf{H}\right)^{-1}\mathbf{H}^{\mathrm{H}}$ is a ZF filter and $\tilde{\mathbf{v}} = \mathbf{H}^{\dagger}\mathbf{v}$ is a transformed noise vector whose covariance matrix is

$$\mathbf{Q}_{\tilde{\mathbf{v}}} = \sigma_v^2\left(\mathbf{H}^{\mathrm{H}}\mathbf{H}\right)^{-1}. \tag{2.25}$$

At the output of the ZF-based detector, the LLR for the nth coded bit of the jth transmit antenna can be obtained using [63] as

$$L(c_{j,n}) = \frac{1}{\sigma_j^2}\left(\min_{x\in\mathcal{A}_n^0} |\tilde{x}_j - x|^2 - \min_{x\in\mathcal{A}_n^1} |\tilde{x}_j - x|^2\right), \qquad j = 1, 2, \ldots, N_t, \tag{2.26}$$

where $\tilde{x}_j = (\tilde{\mathbf{x}}_{\mathrm{ZF}})_j$ is the soft ZF estimate of the jth transmitted symbol and $\sigma_j^2 = (\mathbf{Q}_{\tilde{\mathbf{v}}})_{j,j}$ denotes the noise variance at the output of the ZF filter. Here $\mathcal{A}_n^1$ and $\mathcal{A}_n^0$ denote the sets of scalar symbols in $\mathcal{A}$ whose nth bit equals 1 and 0, respectively.

In the case of an MMSE detector the received vector is passed through a MMSE filter, which results in [63]

$$\tilde{\mathbf{x}}_{MMSE} = \left(\mathbf{H}^{\mathrm{H}}\mathbf{H} + \sigma_v^2\mathbf{I}\right)^{-1}\mathbf{H}^{\mathrm{H}}\mathbf{y} = \mathbf{D}\tilde{\mathbf{x}}_{ZF}, \tag{2.27}$$

where $\mathbf{D} = (\mathbf{I} + \mathbf{Q}_{\tilde{\mathbf{v}}})^{-1} = \left[\mathbf{I} + \sigma_v^2\left(\mathbf{H}^{\mathrm{H}}\mathbf{H}\right)^{-1}\right]^{-1}$. The LLR of $c_{j,n}$ is obtained as

$$L(c_{j,n}) = \frac{1}{\sigma_j^2}\left(\min_{x\in\mathcal{A}_n^0} |\tilde{x}_j - x|^2 - \min_{x\in\mathcal{A}_n^1} |\tilde{x}_j - x|^2\right), \qquad j = 1, 2, \ldots, N_t, \tag{2.28}$$

Here, $\tilde{x}_j = \frac{(\tilde{\mathbf{x}}_{\mathrm{MMSE}})_j}{D_{j,j}}$ denotes the soft MMSE estimate of the jth transmitted symbol and $\sigma_j^2 = \frac{1-D_{j,j}}{D_{j,j}}$ is the noise variance with $D_{j,j}$ representing the (j,j)th diagonal element of $\mathbf{D}$.

2.4 Retransmission Techniques

This section provides a brief overview of various retransmission protocols with and without forward error correction (FEC) codes, e.g. turbo [66, 67] and low-density parity-check (LDPC) [68, 69] codes etc.

2.4.1 ARQ Protocols

Error control techniques such as automatic repeat request (ARQ) can be applied to provide reliable transmission over wired as well as wireless links based on feedback

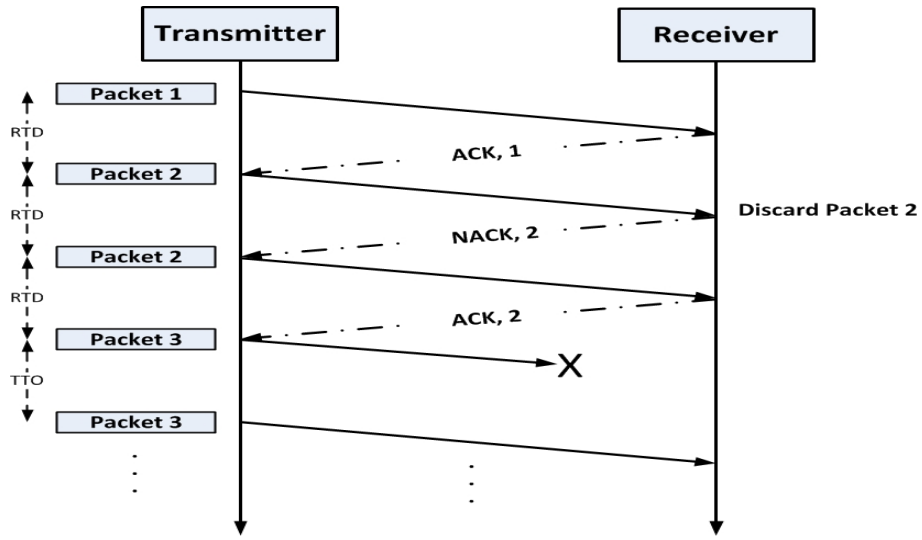


Figure 2.4: Stop-and-wait ARQ scheme.

messages and retransmissions [6]. The feedback channel is used by the receiver to indicate the success or failure of the decoded packet. There are three basic ARQ schemes, namely stop-and-wait (SW), go-back-N (GBN) and selective-repeat (SR) [7]. These ARQ schemes are briefly outlined in the following.

Stop-and-Wait

In the SW ARQ scheme, the transmitter sends a data packet to the receiver and waits for an acknowledgement before transmitting a new packet. The receiver performs a cyclic redundancy check (CRC) on the received packet. If the received packet contains no error, a positive acknowledgment (ACK) is fed back to the transmitter. Otherwise, the receiver discards the erroneously received packet and feeds back a negative acknowledgment (NACK). Upon reception of a NACK message, the transmitter retransmits the corresponding packet. This retransmission process continues until the data packet is correctly received or some pre-defined maximum number of retransmissions is reached.

Figure 2.4 shows the basic procedure of a SW ARQ scheme. The acknowledgment (ACK or NACK) for a packet arrives at the transmitter after a round-trip delay (RTD), which is defined as the time interval between the transmission of a packet and the

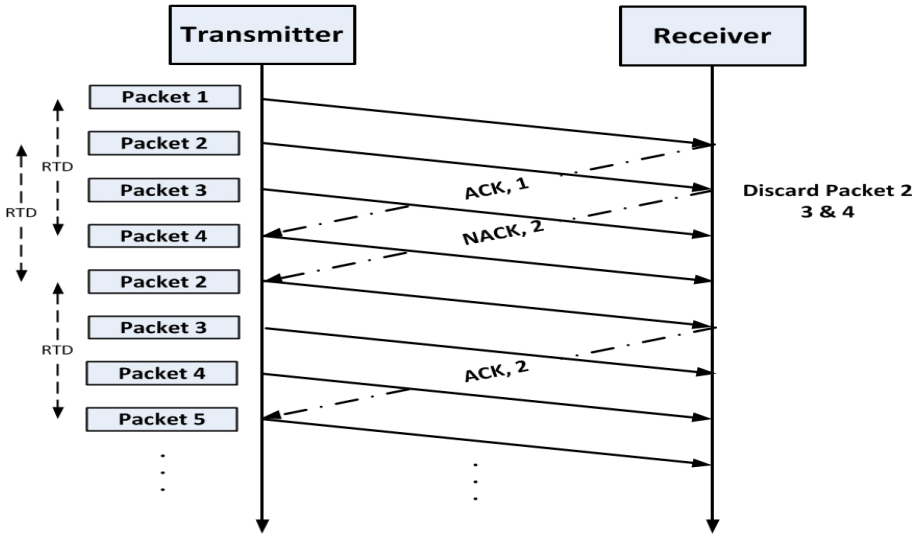


Figure 2.5: Go-back-N ARQ scheme with $N = 3$.

reception of an acknowledgment of that packet [7]. If an acknowledgment is not received by the transmitter within a given time limit, it is defined as a transmission time out (TTO) and the packet is retransmitted.

SW is the simplest form of ARQ schemes with low computational complexity and it operates only over a half-duplex channel [10]. It does not require large buffer size at both ends of the link. However, due to the idle time spent waiting for an acknowledgment of each transmitted packet, it is the most inefficient in terms of throughput compared to more sophisticated ARQ schemes.

Go-Back-N

In the GBN ARQ scheme, the transmitter sends N data packets to the receiver without waiting for an acknowledgment as illustrated in Figure 2.5. Generally the value of N is chosen to be larger than the round-trip delay, so that the transmitter can at least receive an acknowledgment for the first transmitted packet before the end of N transmissions [10]. In case of an erroneously decoded packet (or a lost packet), the receiver discards the erroneously received packet and the subsequent $N - 1$ packets, in spite of whether they are error-free or not [7]. Upon reception of a NACK, the transmitter goes back to

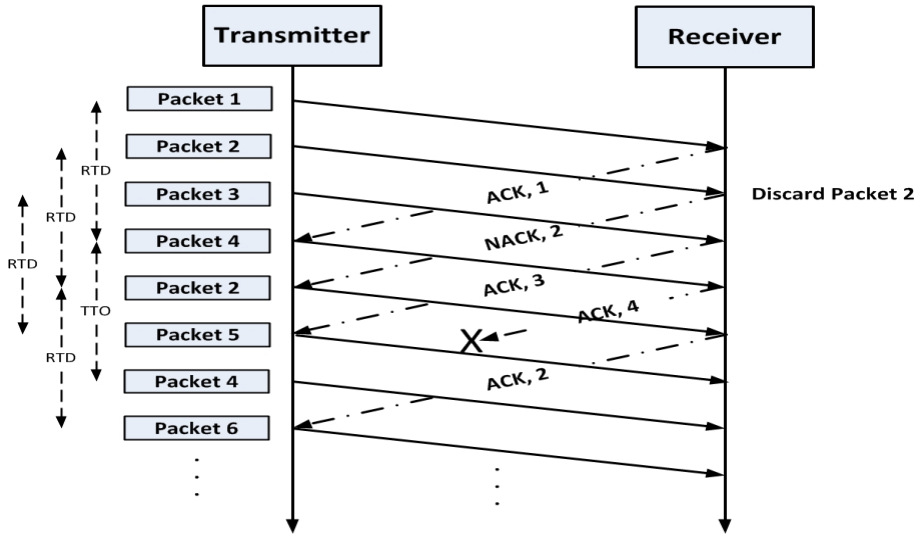


Figure 2.6: Selective-repeat ARQ scheme.

the corresponding packet, stored in the buffer, and retransmits it and the subsequent $N-1$ packets.

One of the major drawbacks of GBN ARQ scheme is that when a packet is erroneously decoded or lost, several packets are retransmitted even if they are successfully decoded at the receiver, resulting in unnecessary retransmissions. To overcome this issue, the SR ARQ scheme is used.

Selective-Repeat

Similar to the GBN scheme, the SR ARQ scheme [7] involves a continuous transmission of data packets. However, the transmitter only retransmits those packets that are erroneously decoded at the receiver as shown in Figure 2.6.

SR scheme is the most efficient ARQ scheme in terms of throughput, but it also requires the largest amount of memory buffer at both ends (i.e. the transmitter and the receiver) of the link [10]. A more detailed comparison of ARQ schemes can be found in [10] and the references therein.

2.4.2 HARQ Protocols

ARQ schemes can be combined with FEC codes to provide reliable transmission over wireless channels that are deeply affected by channel impairments such as fading. The combination of ARQ and FEC techniques results in a so-called hybrid ARQ (HARQ) scheme. Incorporating a FEC code can result in significant reduction in error-rate and accordingly the required number of retransmissions, thus improving the system throughput efficiency.

HARQ schemes can be classified into two categories, namely the type-I HARQ and type-II HARQ. In type-I, when the received packet is unsuccessfully decoded at the receiver, the receiver discards the erroneously decoded packet and asks for a retransmission. The transmitter will then retransmit the same packet. This process continuous until the packet is successfully decoded at the receiver or some maximum retransmission limit is reached. In contrast in type-II, the erroneously decoded packet is not discarded. Instead it is stored in the receiver buffer, and later combined with the following retransmissions. The type-II HARQ protocol can further be divided into two types, namely the packet or Chase combining HARQ (CC-HARQ) [8] and incremental redundancy HARQ (IR-HARQ) [9] also known as code combining. In the following subsections, a brief overview of type-II HARQ protocols is presented.

Chase Combining HARQ

In a Chase combining HARQ (CC-HARQ) scheme [8], the transmitter retransmits the original packet whenever it receives a NACK message[5]. At the receiver, erroneously decoded packets in the previous (re-)transmissions are combined with the current retransmitted packet using the MRC technique. It is known to be the optimum combining method with increased complexity at the receiver [70]. Some other suboptimal combining methods also exist in literature with low complexity compared to MRC, such as SC and EGC [56].

[5]Note that retransmission of the same packet several times, does not bring additional complexity at the transmitter side. However, the cost is a loss in throughput.

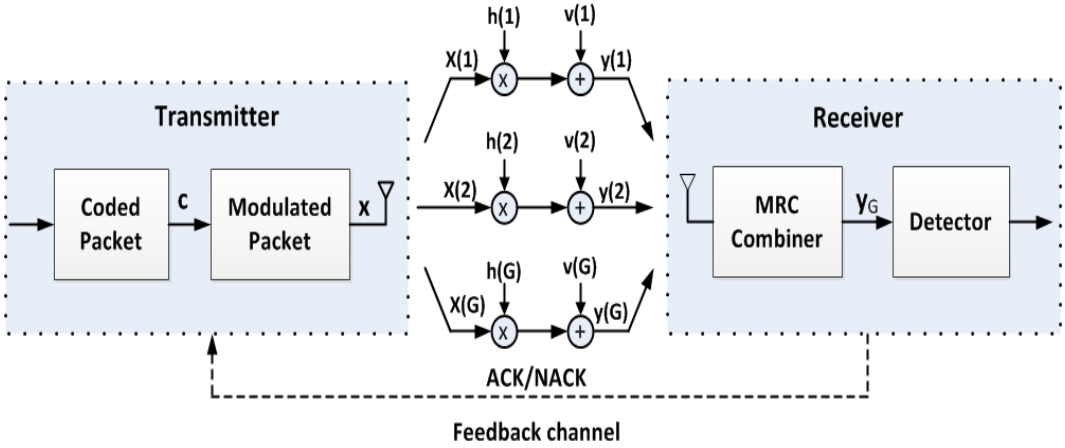


Figure 2.7: Chase combining HARQ (CC-HARQ) scheme.

In order to understand the CC-HARQ scheme, we consider the packet transmission scenario over a SISO channel as shown in Figure 2.7. The channel is assumed to be constant over the packet length L. During the gth (re-)transmission the transmitter (re-)transmits the modulated packet $\mathbf{x}(g)$ containing L complex symbols drawn from an alphabet $\mathcal{A}$ with cardinality $|\mathcal{A}|$. The complex received signal vector during the gth (re-)transmission is written as

$$\mathbf{y}(g) = h(g)\mathbf{x}(g) + \mathbf{v}(g), \qquad g = 1, 2, \ldots, G, \tag{2.29}$$

where $h(g)$ is the Rayleigh fading coefficient and G is the maximum number of transmissions (including $G-1$ retransmissions). Whenever the receiver feeds back a NACK message, the same packet is retransmitted, therefore let $\mathbf{x}(g) = \mathbf{x}$. The instantaneous received SNR during the gth (re-)transmission is given by

$$\gamma(g) = \frac{|h(g)|^2 E_s}{N_0}, \tag{2.30}$$

where E_s is the average energy per symbol and N_0 denotes the one-sided noise spectral density. After G transmissions, the receiver combines the G received copies of the same packet in a single packet. Let $\mathbf{y}_G$ denote the combined packet which is obtained by

weighting each packet with an estimate of its reliability [70], given by

$$\mathbf{y}_G = \sum_{g=1}^{G} \alpha(g)\mathbf{y}(g), \tag{2.31}$$

where $\alpha(g) = h^*(g)/N_0$ is the reliability of each packet. Here $h^*(g)$ denotes the complex conjugate of $h(g)$. This weight combining method is known as MRC. The total SNR with CC-HARQ after G transmissions is given by

$$\begin{aligned} \gamma_G &= \sum_{g=1}^{G} \frac{|h(g)|^2 E_s}{N_0} \\ &= \sum_{g=1}^{G} \gamma(g). \end{aligned} \tag{2.32}$$

Retransmission of the same packet leads to an increased SNR at the receiver due to the MRC technique, thereby increasing the probability of successful decoding. The main drawback of the CC-HARQ scheme is that under excellent channel conditions (high SNR region), meaning that the channel error-rate is low, extra parity bits for error correction are unnecessarily transmitted during each retransmission [10].

Incremental Redundancy HARQ

An incremental redundancy HARQ (IR-HARQ) scheme is proposed to overcome the problem related to CC-HARQ. It is an adaptive HARQ scheme that is most suitable for applications in time-varying channels [10]. The basic idea behind the IR-HARQ scheme is to transmit additional redundancy bits (parity bits), instead of retransmitting the same packet in each retransmission, only when they are needed. To improve decoding reliability the receiver combines all the received versions of the packet.

Figure 2.8 illustrates an example of a conventional IR-HARQ transmission process. The first transmission includes only information bits and a few CRC bits for error detection purposes (similar to a standard ARQ scheme) [10]. In case of unsuccessful decoding, the receiver stores the erroneous packet for combining purposes, and feeds

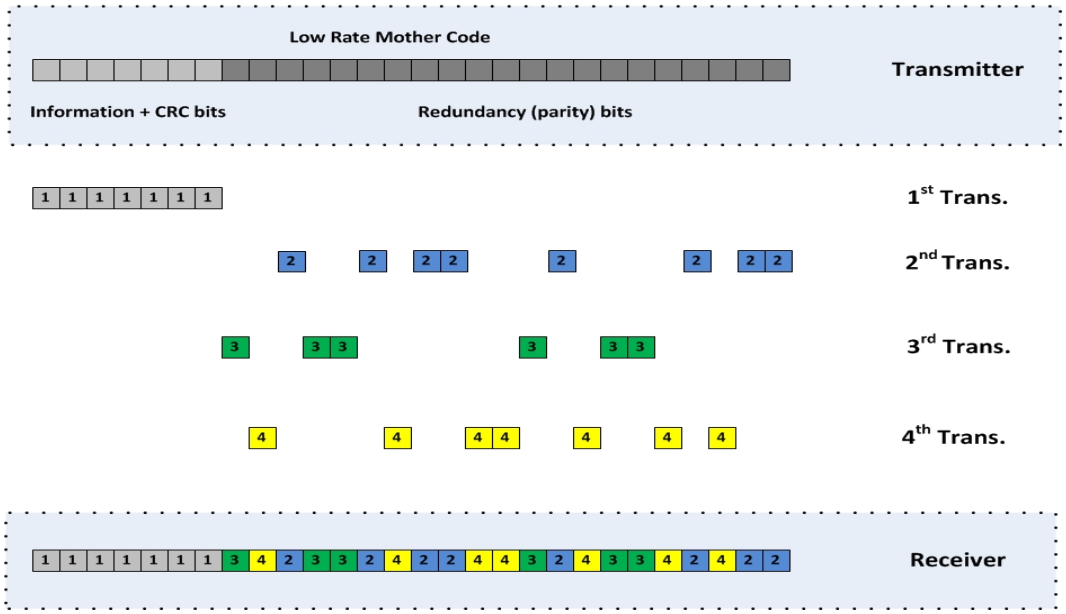


Figure 2.8: Incremental redundancy HARQ (IR-HARQ) scheme.

back a NACK message to the transmitter. The transmitter responds by transmitting only parity bits. This process continues until decoding is successful or some maximum retransmission limit is reached. An IR-HARQ protocol gradually decreases the coding rate, since each retransmission contains a different set of redundancy bits.

Comparing the CC-HARQ protocol with IR-HARQ protocol, CC-HARQ is the simplest HARQ scheme that does not bring additional complexity to the system, since the same packet is retransmitted. On the other hand, IR-HARQ is a more sophisticated HARQ protocol which requires large buffer size and adds more complexity to the system, but it can achieve better performance [10].

2.5 Cooperative Relaying Systems

As mentioned earlier, MIMO technology has been proposed as a powerful means to improve system efficiency and link reliability. In order to fully exploit the spatial diversity gains offered by MIMO systems, the multiple antennas employed at the transmitter and/or at the receiver must be sufficiently spaced, typically one half of the wavelength

to reduce correlation. Unfortunately, implementation of more than one antenna may not always be possible in some applications, such as wireless sensor networks and ad hoc networks, mainly due to size and power constraints. For this reason, the concept of cooperative relaying was proposed, where relay node(s) forward the information of the source node(s) to the destination(s) in order to enhance the network coverage and link reliability. Section 2.5.1 briefly describes the classical relay channel model and its extension to multi-node relaying systems. Signal processing techniques adopted at the relay node(s) are outlined in Section 2.5.2.

2.5.1 Wireless Relay Channel

The key idea of cooperative relaying was originally proposed by van der Meulen in [29, 30]. In these works, a three-node relay channel was introduced which consists of a source, a relay and a destination as illustrated in Figure 1.1. Later, Cover and El Gamal in [31], further investigated the relay channel and derived upper and lower bounds on its capacity. In [31] it was assumed that the relay node operates in full-duplex mode, which means that the relay node transmits and receives simultaneously. However, in current practical systems it may be unrealistic for relays to operate in full-duplex mode [71]. Therefore, throughout the thesis a half-duplex setup is assumed, i.e., the relay cannot transmit and receive at the same time. Furthermore, only an AWGN channel was considered in [29, 30, 31] to analyze the capacity of the relay channel.

In this section, we briefly describe a basic three-node wireless relay system for Rayleigh frequency-flat fading channels, which is shown in Figure 2.9. All three nodes are assumed to have only one antenna. In general, communication in a relay network occurs in two phases, namely the broadcast phase (solid lines) and the relay phase (dashed lines).

In the broadcast phase, the source node transmits its information to the destination. Due to the broadcast nature of the wireless channel, the transmitted information is also received at the relay node. The received signals at the relay and destination, denoted

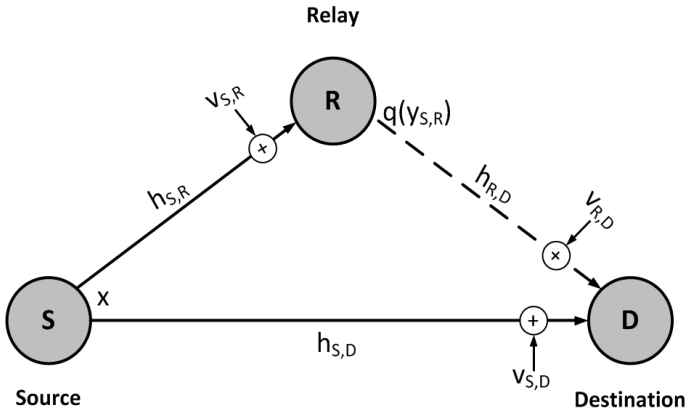


Figure 2.9: Three-node wireless relay network, with a source (S), a relay (R), and a destination (D).

by $y_{S,R}$ and $y_{S,D}$, respectively, can be written as

$$y_{S,R} = h_{S,R}x + v_{S,R}, \tag{2.33}$$

$$y_{S,D} = h_{S,D}x + v_{S,D}, \tag{2.34}$$

where x is the complex transmitted symbol with average energy $\mathbb{E}\{|x|^2\} = 1$, where $\mathbb{E}\{\cdot\}$ denotes the expectation operator. In (2.33) and (2.34), $h_{S,R}$ and $h_{S,D}$ are the channel coefficients between the source and the relay and destination, respectively. These are modeled as independent complex Gaussian random variables with zero-mean and unit variance. Furthermore, $v_{S,R}$ and $v_{S,D}$ are the additive noises at the relay and destination, respectively, which are assumed to be i.i.d. with zero-mean and variance N_0.

In the relay phase, the relay forwards the processed version of the source's information to the destination. The received signal at the destination is then given by

$$y_{R,D} = h_{S,D}q(y_{S,R}) + v_{R,D}, \tag{2.35}$$

where the function $q(\cdot)$ depends on the signal processing technique adopted at the relay node [72]. Some of the widely studied relaying protocols are discussed in Section 2.5.2.

The destination then combines the signals from the source and relay nodes in order to decode the information.

The instantaneous and average received SNRs for the link between transmitting node a (S or R) and receiving node b (R or D) are given respectively by $\gamma_{a,b} = |h_{a,b}|^2 E_a/N_0$ and $\bar{\gamma}_{a,b} = \mathbb{E}\{\gamma_{a,b}\}$, where E_a denotes the average energy per symbol transmitted from node a.

Extension to Multi-node Systems

The three-node relay channel discussed above can be extended to large network configurations, such as cooperative multiple access (uplink) channels and broadcast (downlink) channels (e.g. [73, 74, 75]). We focus only on the uplink cooperative network, where multiple relay nodes[6] assist in forwarding the information of multiple source nodes to a common destination (e.g. base station), as illustrated in Figure 2.10. Signals from different source nodes are transmitted through the same cooperative channel to improve bandwidth (spectral) efficiency. Common multiple access techniques are time division multiple access (TDMA), frequency division multiple access (FDMA), code division multiple access (CDMA) and space division multiple access (SDMA).

In TDMA or FDMA, each transmitting node is assigned an orthogonal time or frequency channel to conduct its transmission, thus avoiding the multiple access interference (MAI). In case of two-hop transmission as shown in Figure 2.10, the orthogonal channel can be further divided in two sub-channels. More specifically, for TDMA transmissions, each time slot is divided into two sub-slots, used for the broadcast phase (solid lines) and relay phase (dashed lines) transmissions performed by the source and the relay, respectively [76]. The same is the case for FDMA transmissions, where sub-channel allocation is in frequency bands rather than in time slots. In CDMA[7] or SDMA, source nodes simultaneously transmit over different codes or spatial dimensions. As a result,

[6] The relay nodes are assumed to be operating in half-duplex mode.

[7] CDMA technique can be implemented for cooperative systems in orthogonal fashion by using different spreading codes for each user to avoid interference. However, due to lack of perfect synchronizing at the receiver, the requirement for orthogonality is difficult to satisfy in practical systems [76].

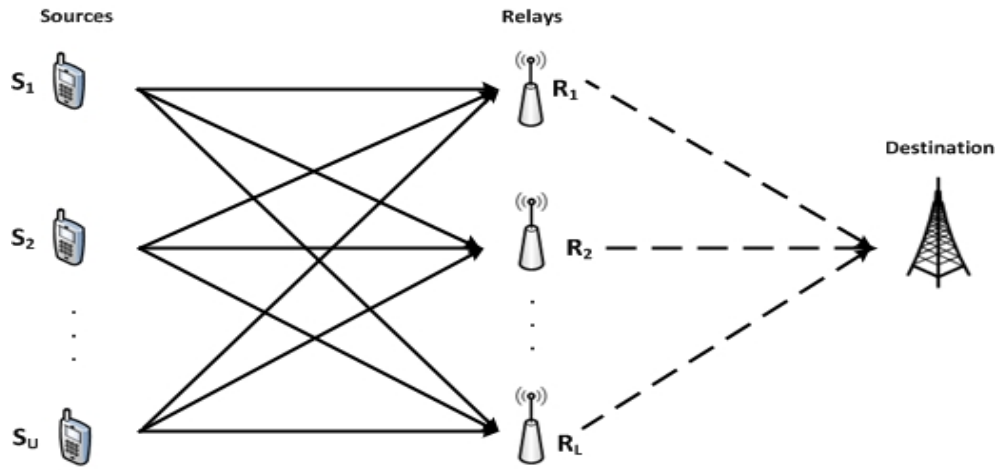


Figure 2.10: Illustration of multi-source multi-relay network.

MAI is introduced at the receiver (relays and destination). To combat MAI, MUD techniques can be employed at the receiver to jointly detect the transmitted information from the multiple source nodes.

2.5.2 Relaying Protocols

Numerous relaying protocols have been proposed in the literature, among which amplify-and-forward (AF) and decode-and-forward (DF) are the simplest and most widely adopted [26, 32]. In this section, these protocols are discussed for a single relay node case, but can be applied to multi-relay scenarios.

Amplify-and-Forward

Amplify-and-forward is the simplest relaying protocol, which was proposed and analyzed by Laneman in [77]. It is often called a non-regenerative protocol. As the name implies, the relay node simply amplifies the received signal from the source and forwards it to the destination during the relay phase as shown in Figure 2.11(a). A major shortcoming of this protocol is that amplifying the received signal will also amplify the additive noise at the relay. For the AF protocol, the function $q(\cdot)$ as described in (2.35)

can be expressed as

$$q(y_{S,R}) = \beta y_{S,R}, \tag{2.36}$$

where β is the amplifying factor given by

$$\beta = \sqrt{\frac{E_R}{|h_{S,R}|^2 E_S + N_0}}. \tag{2.37}$$

According to (2.37), the received signal at the destination from the relay can be written as

$$y_{R,D} = \sqrt{\frac{E_R}{|h_{S,R}|^2 E_S + N_0}} h_{R,D} y_{S,R} + v_{R,D}, \tag{2.38}$$

where $h_{R,D}$ is the fading coefficient between the relay and the destination, and $v_{R,D}$ is the additive noise at the destination. From (2.33) and (2.38), the received signal $y_{R,D}$ in this case is

$$y_{R,D} = \sqrt{\frac{E_R}{|h_{S,R}|^2 E_S + N_0}} h_{R,D} h_{S,R} x + \tilde{v}_{R,D}, \tag{2.39}$$

where $\tilde{v}_{R,D}$ is the effective additive Gaussian noise with zero-mean and variance [72]

$$\tilde{N}_0 = \left(\frac{|h_{R,D}|^2 E_R}{|h_{S,R}|^2 E_S + N_0} + 1 \right) N_0. \tag{2.40}$$

A MRC technique is applied at the destination to combine the received signals $y_{S,D}$ and $y_{R,D}$. With perfect knowledge of the channel coefficients $h_{S,D}$, $h_{S,R}$ and $h_{R,D}$, available at the receiving nodes, the combined received signal y_D at the output of the MRC can be expressed as

$$y_D = \alpha_{S,D} y_{S,D} + \alpha_{R,D} y_{R,D}, \tag{2.41}$$

where $\alpha_{S,D} = h^*_{S,D}/N_0$ and $\alpha_{R,D} = \beta h^*_{S,D} h^*_{R,D}/\tilde{N}_0$ are the combining coefficients. For the uncoded transmissions, the destination computes the hard estimate of x by using ML detection as:

$$\hat{x} = \arg \min_{x \in \mathcal{A}} \| y_D - \alpha_{S,D} h_{S,D} x - \alpha_{R,D} h_{R,D} x \|^2. \tag{2.42}$$

However, for coded transmissions, the destination employs soft-output ML detection algorithm to calculate the corresponding soft estimates on x. The resulting soft estimates are then passed to the channel decoder to make hard decisions. Other suboptimal detection techniques for uncoded and coded transmissions can also be used, and some of them are discussed in Section 2.3.

The total instantaneous SNR at the output of MRC at the destination is equal to the sum of the instantaneous received SNRs from both transmission phases. It is expressed as [78]

$$\gamma_{\text{tot}} = \gamma_{S,D} + \frac{\gamma_{S,R}\gamma_{R,D}}{\gamma_{S,R} + \gamma_{R,D} + 1}, \tag{2.43}$$

where $\gamma_{S,D} = |h_{S,D}|^2 E_S/N_0$, $\gamma_{S,R} = |h_{S,R}|^2 E_S/N_0$, $\gamma_{R,D} = |h_{R,D}|^2 E_R/N_0$ are the instantaneous received SNRs at the destination from the source, at the relay from the source, and at the destination from the relay, respectively.

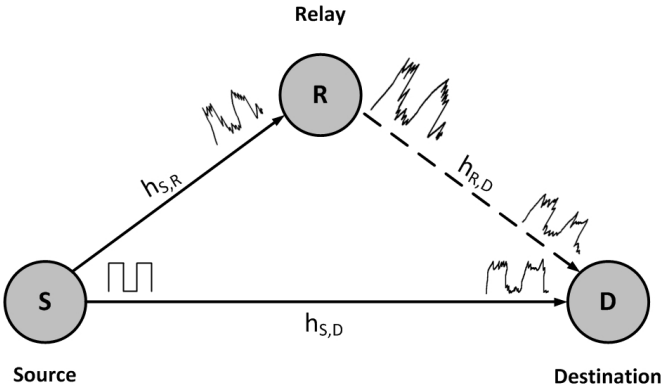
Decode-and-Forward

In decode-and-forward protocols (also refereed as re-generative protocols), the relay node decodes the received signal from the source node during the broadcast phase, and forwards it either as is or the re-encoded signal to the destination during the relay phase as shown in Figure 2.11(b). This type of relaying scheme was first introduced by Cover and El Gamal in [31]. For uncoded DF relaying, let the function $q(y_{R,D}) = \hat{x}$, where $\hat{x}$ is the estimated symbol at the relay node. The relay node then forwards $\hat{x}$ as it is to the destination. The received signal $y_{R,D}$ at the destination is thus given by

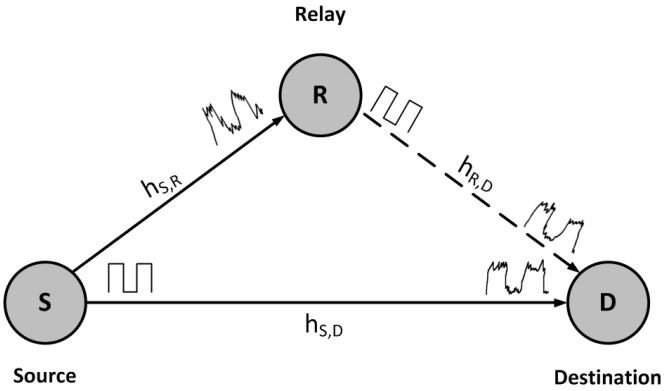
$$y_{R,D} = h_{R,D}\hat{x} + v_{R,D}. \tag{2.44}$$

For coded DF relaying, let the function $q(y_{R,D}) = x_R$, where x_r is the re-encoded information at the relay node. The received signal $y_{R,D}$ at the destination in this case is

$$y_{R,D} = h_{R,D}x_R + v_{R,D}. \tag{2.45}$$



(a) Amplify-and-Forward Protocol (AF).



(b) Decode-and-Forward Protocol (DF).

Figure 2.11: Comparison of amplify-and-forward (AF) and decode-and-forward (DF) relaying protocols.

The combined received signal y_D of the MRC output at the destination can be expressed as (2.41). However, in the case of DF relaying protocols the combining coefficients are defined as $\alpha_{S,D} = h^*_{S,D}/N_0$ and $\alpha_{R,D} = h^*_{R,D}/N_0$.

The DF relaying protocol provides superior performance compared to the AF relaying protocol when the source-relay channel quality is good, as successful decoding and then re-encoding can completely avoid noise amplification at the relay node [79]. However, when the source-relay channel quality is poor, the AF protocol can provide better performance, because forwarding erroneously decoded signals (at the relay) to the destination can cause error propagation and consequently severe degradation in system performance. In order to overcome the problem of error propagation, a selective

DF [80, 81] relaying scheme was proposed. Unlike the conventional DF relaying, when the received signals at the relay contains errors (according to CRC), the relay node remains silent. In this thesis selective DF relaying protocol is adopted for cooperative HARQ systems. For convenience, the selective DF relaying considered in Chapter 4 and 5 will simply be called DF relaying in those chapters.

2.6 Chapter Conclusions

This chapter introduced several concepts and assumptions that will be used in the subsequent chapters. First, wireless channel characteristic and statistical channel models were summarized. Thereafter, system models for point-to-point MIMO and MU-MIMO systems were presented. Numerous multi-signal detection algorithms for uncoded and coded transmissions were also discussed. Furthermore, an overview of retransmission protocols with and without FEC codes was presented. Finally, the concept of cooperative relaying systems was introduced, which included an overview of three-node relay channel and various relaying protocols adopted at the relay node. Extension to multi-source multi-relay cooperative system was also summarized.

Chapter 3

Non-Cooperative Overloaded Systems

This chapter focuses on studying new signal processing techniques for uplink multi-user non-cooperative systems under overloaded conditions, that is, where there are fewer receive antennas (N_r) than transmitting users/antennas (N_t).

3.1 Introduction

In many communication environments, the number of users varies with time. As a result, a system may fluctuate among the three loading conditions, of being underloaded ($N_r < N_t$), critically loaded ($N_r = N_t$) and overloaded ($N_r > N_t$). Generally for most multiple antenna systems, it is assumed that $N_r \geq N_t$. Under such conditions, various optimal and suboptimal multi-user detection (MUD) algorithms can easily be employed at the receiver to recover the transmitted signals, with a simple trade-off between complexity and performance. However, MUD for overloaded systems is a challenging task. Co-channel interference (CCI) introduced by multiple users severely degrades the performance of linear MUD algorithms under overloaded conditions compared to critically loaded and underloaded conditions [64]. Suboptimal non-linear interference cancellation (IC) algorithms [14] fail under overloaded conditions, due to the channel matrix singularity [15]. The optimum maximum a posteriori (MAP) and maximum

likelihood (ML) detectors work well under these conditions [17]. Unfortunately, their exponential complexity make them impractical for most applications. Various other suboptimal detection techniques [15, 17, 19, 20] have been proposed for overloaded systems. However, these suboptimal algorithms demand sophisticated implementation and still have quite high computational complexity [21].

Channel augmentation techniques can be applied at the transmitter and/or at the receiver to form multiple virtual antennas, thereby improving the system performance. Rankin *et al.* [23] showed that by repetitively transmitting the same spatial multiplexed signals, the information outage rate of multiple-input multiple-output (MIMO) channels can be improved. More specifically, the authors stacked the received signal vectors of the same transmitted signals to create additional virtual receive antennas. Recently in [24], a linear MIMO hybrid automatic repeat request (HARQ) precoder was designed to optimize the mutual information, and joint HARQ detection was performed by stacking the received vectors from all (re-)transmissions. However, only critically loaded conditions were considered in [24].

In this chapter, HARQ retransmissions with the use of virtual receive antennas [23] and a simple MUD algorithm are combined together resulting in a novel multi-user transmission approach that works well under overloaded conditions. Unlike the work of [24], "stacked"transmissions are used in the formation of virtual receive antennas, which transform an overloaded system into a critically or underloaded system. Furthermore in [24], the same signal vector was transmitted from all antennas during all (re-)transmissions. In our opinion, this wastes resources; users with successfully decoded packets do not want to retransmit their packets in practical systems. To overcome this drawback two new multi process HARQ schemes are proposed. In Scheme-I, users with correctly decoded packets remain idle, whereas users with erroneous packets retransmit their packets in subsequent frames. In Scheme-II, users with correctly decoded packets may transmit new information instead of remaining idle. These schemes allow us to apply linear MUD algorithms without requiring additional antennas or

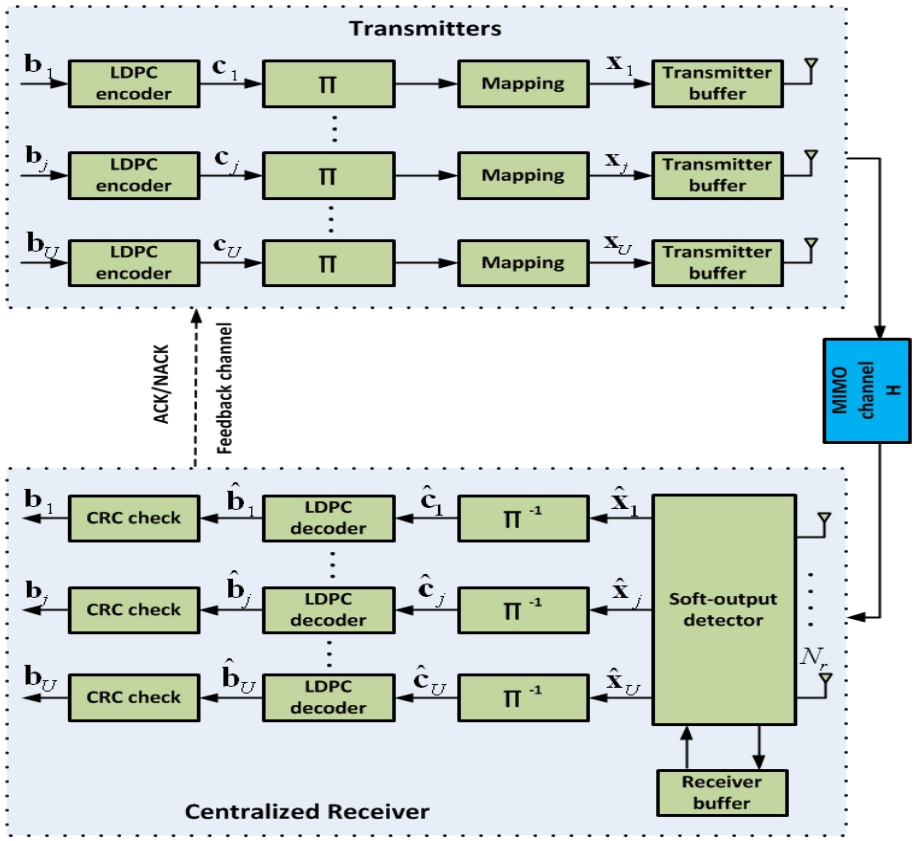


Figure 3.1: Block diagram of U co-channel users/transmitters and a centralized receiver.

hardware chains.

This chapter is organized as follows. In Section 3.2, the system and channel model is described. Section 3.3, presents the proposed HARQ schemes under time-varying loading conditions. In Section 3.4, simulation results are presented. Finally, conclusions are drawn in Section 3.5.

3.2 System Model

The uplink multi-user system model described here is similar to that defined in Section 2.2.2, but is repeated here for convenience. The block diagram of an uplink multi-user system is shown in Figure 3.1, consisting of three main parts: (i) U co-channel users/transmitters each with a single antenna, (ii) the MIMO channel, and (iii) the

centralized receiver with N_r receive antennas.

3.2.1 Transmitters

For each transmitter, a packet $\mathbf{b}_j = (b_{j,1}, \cdots, b_{j,k}, \cdots, b_{j,K})$ containing K bits, which includes both information and cyclic redundancy check (CRC) bits, is encoded by a low-density parity-check (LDPC) encoder of rate $R_c = K/N$, resulting in a codeword $\mathbf{c}_j = (c_{j,1}, \cdots, c_{j,n}, \cdots, c_{j,N})$ of length N. Here $j = 1, 2, \cdots, U$ is the user index. Since each user is assumed to be equipped with a single antenna, the total number of antennas across all users is $N_t = U$. The coded bits are then re-ordered by random bit-interleaver (Π). For a given symbol interval, a group of $\log_2(C)$ interleaved coded bits are mapped into a modulated symbol x_j taken from an alphabet $\mathcal{A}$ and transmitted over the wireless channel, where C represents the alphabet size given in (2.5).

3.2.2 MIMO Channel

The complex received signal vector $\mathbf{y} \in \mathbb{C}^{N_r \times 1}$ for a symbol interval is expressed as [60]

$$\begin{aligned} \mathbf{y} &= \sum_{j=1}^{N_t} \mathbf{h}_j x_j + \mathbf{v} \\ &= \mathbf{H}\mathbf{x} + \mathbf{v}, \end{aligned} \tag{3.1}$$

where $\mathbf{h}_j$ is the jth column of the channel matrix $\mathbf{H}$ of size $N_r \times N_t$, $\mathbf{x} \in \mathbb{C}^{N_t \times 1}$ is the overall transmitted vector and $\mathbf{v} \in \mathbb{C}^{N_r \times 1}$ is a complex additive white Gaussian noise (AWGN) vector whose components each have zero-mean and variance σ_v^2 . The total number of symbol intervals for each transmission round is equal to the packet length L. The received signal matrix after L symbol intervals is then given by [60]

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{V}, \tag{3.2}$$

where $\mathbf{Y} \in \mathbb{C}^{N_r \times L}$, $\mathbf{X} \in \mathbb{C}^{N_t \times L}$ and $\mathbf{V} \in \mathbb{C}^{N_r \times L}$. The channel is assumed to exhibit quasi-static, frequency-flat Rayleigh fading, meaning the channel gains are constant over one packet transmission and vary independently between packets. The channel coefficients $h_{i,j}$ between receive antenna i and transmit antenna j are modeled as independent identically distributed (i.i.d) complex Gaussian random variables with zero-mean and unit variance. Furthermore, it is assumed that perfect channel state information (CSI) is available at the receiver.

3.2.3 Centralized Receiver

The receiver employs soft-output MUD, either optimal or suboptimal linear detection (described in Section 2.3.2) to reduce CCI due to multiple users. The resulting soft detector estimates are passed to bit-deinterleavers and then to the channel decoders of each user. It is assumed that a CRC decoder performs perfect error detection at the receiver. If a packet is found to be error-free after passing through the LDPC decoder a positive acknowledgement (ACK) is sent to the corresponding transmitter, otherwise a negative acknowledgement (NACK) is sent requesting retransmission of the same packet. We assume that the feedback channel is error free.

3.3 Proposed Schemes

In this section, the proposed HARQ schemes for variable loading multi-user systems are presented. The basic idea behind the proposed schemes is to treat the HARQ retransmission information as if it was arriving to an additional (virtual) N_r receive antennas [23]. From (3.2) the received signal matrix at the gth (re-)transmission can be written as

$$\mathbf{Y}(g) = \mathbf{H}(g)\mathbf{X}(g) + \mathbf{V}(g), \qquad g = 1, 2, \ldots G, \tag{3.3}$$

where G denotes the maximum number of transmissions (including $G-1$ retransmissions). Retransmissions are assumed to be delayed long enough to result in independent

channel conditions for each transmission from a given user. The stacked received vectors after G transmissions can be written as [24]

$$\mathbf{R}_G = \begin{bmatrix} \mathbf{Y}(1) \\ \mathbf{Y}(2) \\ \vdots \\ \mathbf{Y}(G) \end{bmatrix} = \begin{bmatrix} \mathbf{H}(1) \\ \mathbf{H}(2) \\ \vdots \\ \mathbf{H}(G) \end{bmatrix} \mathbf{X} + \begin{bmatrix} \mathbf{V}(1) \\ \mathbf{V}(2) \\ \vdots \\ \mathbf{V}(G) \end{bmatrix}, \tag{3.4}$$

where the components of $\big(\mathbf{H}(1), \mathbf{H}(2), \ldots, \mathbf{H}(G)\big)$ and $\big(\mathbf{V}(1), \mathbf{V}(2), \ldots, \mathbf{V}(G)\big)$ are independent. One of the major drawbacks of (3.4) is its wasted resources, since the same $\mathbf{X}$ is retransmitted from all users during all retransmissions. In practical systems, a user whose packet is successfully decoded at the receiver does not want to retransmit its packet in subsequent frames. To overcome this drawback (i.e. to avoid retransmissions from users with successfully decoded packets), two alternative schemes are described below.

3.3.1 Scheme-I

In Scheme-I, users with correctly decoded packets remain idle, while users with erroneous packets retransmit their packets in the next time interval. Scheme-I results in improved signal-to-interference-plus-noise ratio (SINR) for later transmissions and hence improves performance. Let $\mathcal{S}$ and $\mathcal{E}$ denotes the set of indices of user data packets that are successfully and erroneously decoded at the receiver, respectively. The steps involved in scheme-I are as follows:

1. During the first transmission, the receiver performs CRC on the decoded packets and accordingly feeds back an ACK or a NACK to each user. The contributions of correctly received users are removed from the received signal vector. This is done by re-encoding, re-mapping the estimates of the successfully decoded packets, and then subtracting them from the received matrix. The modified (interference

cancelled) received signal matrix for the erroneous users is then given by

$$\bar{\mathbf{Y}}(1) = \mathbf{Y}(1) - \sum_{q \in \mathcal{S}} \mathbf{h}_q(1)\hat{\mathbf{x}}_q(1), \tag{3.5}$$

where index q refers to the qth user whose packet is successfully decoded and $\hat{\mathbf{x}}_q(1)$ denotes the decision made for the transmitted symbol vector $\mathbf{x}_q(1)$.

2. During the second transmission, each user on receiving a NACK responds by retransmitting its packet, while it remains silent after receiving an ACK in order to reduce interference. The received signal matrix is then given by

$$\mathbf{Y}(2) = \mathbf{H}(2)\mathbf{X}(2) + \mathbf{V}(2), \tag{3.6}$$

where the column(s) and row(s) of the channel matrix $\mathbf{H}(2)$ and the transmitted symbol matrix $\mathbf{X}(2)$, respectively correspond only to the retransmitting user(s).

3. The received signal matrix $\mathbf{Y}(2)$ is stacked with $\bar{\mathbf{Y}}(1)$ to virtually create $(2N_r)$ receive antennas, resulting in

$$\mathbf{R}_2 = \begin{bmatrix} \bar{\mathbf{Y}}(1) \\ \mathbf{Y}(2) \end{bmatrix}. \tag{3.7}$$

4. Soft data estimates using the stacked received vectors are calculated by the detector and passed to the channel decoders. If any user's retransmitted packet is still decoded in error, the receiver again requests a retransmission. However, users with successfully decoded retransmitted packets remain idle and their contribution from previous transmissions is cancelled. This retransmission is treated as creating an additional N_r virtual receive antennas and stacked with the previous transmissions. MUD is applied to the resulting $gN_r \times L$ matrix, $\mathbf{R}_g = \left[\bar{\mathbf{Y}}(1); \ldots; \bar{\mathbf{Y}}(f); \mathbf{Y}(g)\right]$, with $\bar{\mathbf{Y}}(f) = \mathbf{Y}(f) - \sum_{q \in \mathcal{S}} \mathbf{h}_q(f)\hat{\mathbf{x}}_q(f)$, $1 \leq f \leq g-1$, to provide reliable soft estimates as inputs to the decoders. This process continues

until all packets are successfully decoded or the maximum number of transmission, G, is reached, which is the same for all users. If a packet is still in error after G transmissions, it is declared a dropped packet. Table 3.1 summarizes the proposed HARQ scheme-I algorithm.

	Scheme-I
1:	**Initialize**: $g = 1$, $f = 1$, $\mathcal{S} = \emptyset$, and $\mathcal{E} = \emptyset$.
2:	Perform soft-output ML or MMSE MUD on the received matrix $\mathbf{Y}(1)$ using (2.23) or (2.28), respectively, in order to calculate the LLRs for each user's packet.
3:	Calculate the CRC for each user's decoded packet.
4:	Update $\mathcal{S}$ and $\mathcal{E}$.
5:	Feed back an ACK for each user $q \in \mathcal{S}$, and a NACK for each user $e \in \mathcal{E}$, where $q \neq e$.
6:	**if** $\mathcal{E} \neq \emptyset$ **then**
7:	Update $g = g + 1$.
8:	**repeat**
9:	Remove the contribution of successfully decoded packets from the received signal matrix $\mathbf{Y}(f)$ to obtain $\bar{\mathbf{Y}}(f) = \mathbf{Y}(f) - \sum_{q \in \mathcal{S}} \mathbf{h}_q(f)\hat{\mathbf{x}}_q(f).$
10:	User(s) belonging to the set $\mathcal{E}$ will retransmit during the gth transmission, while all successfully decoded user(s) belonging to the set $\mathcal{S}$ will remain silent.
11:	Obtain the received signal matrix, $\mathbf{Y}(g)$, for the gth transmission and then stack it with the previous post-cancelled received matrices.
12:	Perform MUD on the resulting matrix, $\mathbf{R}_g = \left[\bar{\mathbf{Y}}(1); \ldots; \bar{\mathbf{Y}}(f); \mathbf{Y}(g)\right]$, to obtain the LLRs and calculate the CRC for each user's decoded packet belonging to the set $\mathcal{E}$.
13:	Update $\mathcal{S}$ and $\mathcal{E}$.
14:	Feed back an ACK for each user $q \in \mathcal{S}$ and a NACK for each user $e \in \mathcal{E}$.
15:	**if** $g = G$ and $\mathcal{E} \neq \emptyset$ **then**
16:	Declare a dropped packet for each user $e \in \mathcal{E}$.
17:	**end if**
18:	Update $f = f + 1$ and $g = g + 1$.
19:	**until** $g > G$ or $\mathcal{E} = \emptyset$
20:	**end if**
21:	All users will transmit new packets.

Table 3.1: Proposed HARQ Scheme-I algorithm.

3.3.2 Scheme-II

This scheme allows the transmission of new information from the correctly decoded users instead of their remaining idle. The details of scheme-II are as follows:

1. After the first transmission CRC, the receiver updates $\mathcal{S}$ and $\mathcal{E}$, and feeds back an ACK for each user $q \in \mathcal{S}$, and a NACK for each user $e \in \mathcal{E}$. The channel matrix $\mathbf{H}(1)$ is modified by zeroing out the columns representing the correctly decoded users, $\bar{\mathbf{H}}(1) = \left[\bar{\mathbf{h}}_1(1) \cdots \bar{\mathbf{h}}_j(1) \cdots \bar{\mathbf{h}}_{N_t}(1)\right]$. Let $\bar{\mathbf{h}}_j(1) = \mathbf{0}$, $\forall j \in \mathcal{E}$ and $\bar{\mathbf{h}}_j(1) = \mathbf{h}_j(1)$, $\forall j \in \mathcal{S}$, where the index $j = 1, 2, \ldots, N_t$, refers to the jth column in $\bar{\mathbf{H}}(1)$ and $\mathbf{H}(1)$. The modified received signal matrix is then given by

$$\bar{\mathbf{Y}}(1) = \bar{\mathbf{H}}(1)\mathbf{X}(1) + \mathbf{V}(1). \tag{3.8}$$

2. During the second time interval, user(s) belonging to the set $\mathcal{E}$ will retransmit the same packets, while the remaining user(s) may transmit new packets. Therefore, the received signal matrix $\mathbf{Y}(2)$ may contain both new and retransmitted information. Stacking both received signal matrices in virtual sense creates $(2N_r)$ receive antennas, resulting in

$$\mathbf{R}_2 = \begin{bmatrix} \bar{\mathbf{Y}}(1) \\ \mathbf{Y}(2) \end{bmatrix}. \tag{3.9}$$

3. This process continues until the retransmitted and new transmitted packets are successfully decoded or the maximum number of transmissions, G_j, $\forall j$, is reached. If the packet of user $e \in \mathcal{E}$ is still in error after G_e transmissions, it is declared a dropped packet. Table 3.2 summarizes the proposed HARQ scheme-II algorithm.

	Scheme-II
1:	**Initialize**: $g_j = 1$, $m = 1$, $f = 1$, $\mathcal{S} = \emptyset$, and $\mathcal{E} = \emptyset$.
2:	Perform soft-output ML or MMSE MUD on the received matrix $\mathbf{Y}(1)$ using (2.23) or (2.28), respectively, in order to calculate the LLRs for each user's packet.
3:	Calculate the CRC for user's decoded packet.
4:	Update $\mathcal{S}$ and $\mathcal{E}$.
5:	Feed back an ACK for each user $q \in \mathcal{S}$, and a NACK for each user $e \in \mathcal{E}$, where $q \neq e$.
6:	**if** $\mathcal{E} \neq \emptyset$ **then**
7:	Update $m = m + 1$.
8:	**repeat**
9:	For each user $j \in \mathcal{E}$ set $\bar{\mathbf{h}}_j(f) = \mathbf{0}$ and update $g_j = g_j + 1$.
10:	For each user $j \in \mathcal{S}$ set $\bar{\mathbf{h}}_j(f) = \mathbf{h}_j(f)$ and set $g_j = 1$.
11:	Each user belonging to the set $\mathcal{E}$ will retransmit its packet, while each successfully decoded user belonging to the set $\mathcal{S}$ will transmit a new packet in the mth transmission.
12:	Obtain the received signal matrix, $\mathbf{Y}(m)$, and then stack it with the previous modified received matrices, $\big(\bar{\mathbf{Y}}(1), \ldots, \bar{\mathbf{Y}}(f)\big)$, where $\bar{\mathbf{Y}}(f) = \bar{\mathbf{H}}(f)\mathbf{X}(f) + \mathbf{V}(f)$ and $\bar{\mathbf{H}}(f) = \big[\bar{\mathbf{h}}_1(f) \cdots \bar{\mathbf{h}}_j(f) \cdots \bar{\mathbf{h}}_{N_t}(f)\big]$.
13:	Perform MUD on the resulting matrix, $\mathbf{R}_m = \big[\bar{\mathbf{Y}}(1); \ldots; \bar{\mathbf{Y}}(f); \mathbf{Y}(m)\big]$, to obtain the LLRs and calculate the CRC for user's decoded packet.
14:	Update $\mathcal{S}$ and $\mathcal{E}$.
15:	Feed back an ACK for each user $q \in \mathcal{S}$ and a NACK for each user $e \in \mathcal{E}$.
16:	**for** $j = 1, 2, \ldots, N_t$ **do**
17:	**if** $g_j = G$ and $j \in \mathcal{E}$ **then**
18:	Declare a dropped packet for user j and set $g_j = 0$.
19:	**end if**
20:	**end for**
21:	Update $f = f + 1$ and $m = m + 1$.
22:	**until** $\mathcal{E} = \emptyset$
23:	**end if**
24:	All users will transmit new packets.

Table 3.2: Proposed HARQ Scheme-II algorithm.

3.4 Simulation Results

The simulation results for two ($N_t = 2$) and four ($N_t = 4$) user uplink systems are presented in this section. Each packet contains $K = 576$ information bits (including data and CRC bits), and these are encoded by a rate $R_c = 1/2$ LDPC code. Quasi-

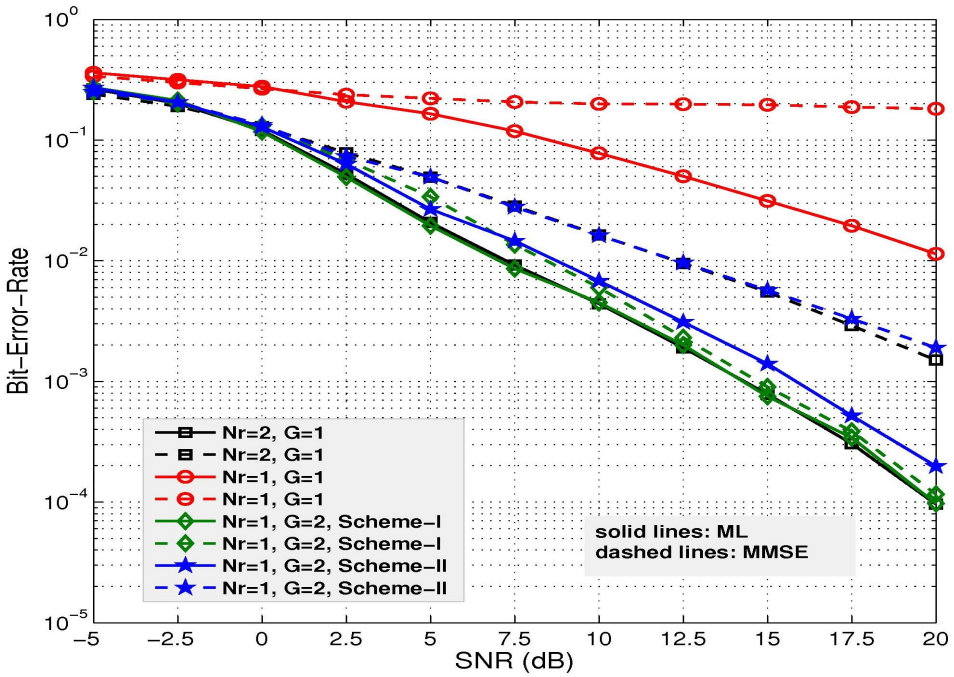


Figure 3.2: BER performance of proposed two user schemes, $N_t = 2$, using soft output ML (solid lines) and linear MMSE MUD (dashed lines).

static Rayleigh frequency-flat fading channel and 4-quadrature amplitude modulation (QAM) modulation are considered. Performance is evaluated in terms of bit-error-rate (BER), dropped packet-rate ($\mathcal{P}_{rate}$) and throughput spectral efficiency (η) as a function of the bit energy-to-noise density ratio E_b/N_0 at each receive antenna, where

$$\frac{E_b}{N_0} = \frac{E_s}{N_0 \log_2 |C| R_c},$$

E_s is the average energy per symbol and N_0 is the one-sided noise spectral density.

The BER performance for an overloaded 2×1 ($N_t = 2, N_r = 1$) and critically loaded 2×2 ($N_t = 2, N_r = 2$) system with optimal ML and suboptimal linear MMSE soft output detectors is presented in Figure 3.2. Clearly a critically loaded system ($N_r = 2$) outperforms an overloaded system ($N_r = 1$) for the same number of transmissions, due to its higher diversity order. However, scheme-I for the overloaded system with two transmissions ($N_r = 1, G = 2, GN_r = 2$) employing ML achieves similar BER performance to the critically loaded case with one transmission ($N_r = 2, G = 1, GN_r = 2$).

There is a slight performance degradation when scheme-II is considered for the 2×1 system with $G = 2$. Furthermore, it is clear from Figure 3.2 that when employing MMSE for a 2×1 overloaded system scheme-I achieves a significant performance gain over scheme-II, due to the fact that scheme-II sends new information during retransmissions thus decreasing the received SINR.

The resulting BER performance for a four user HARQ system using both ML and linear MMSE soft output detectors is presented in Figures 3.3 and 3.4, respectively. As illustrated, a 4×1 overloaded system with no retransmissions ($G = 1$) performs poorly. The proposed schemes allow us to transform overloaded conditions to critically loaded or underloaded conditions by allowing sufficient transmissions. For example, as shown in Figure 3.3 scheme-I for an overloaded 4×1 system with ML MUD achieves critically loaded 4×4 system performance using $G = 4$ transmissions. Similarly, a 4×2 overloaded system using scheme-I requires $G = 2$ transmissions to achieve the same diversity and performance as a 4×4 system that uses $G = 1$.

Figure 3.4 shows that the 4×1 scheme-I system employing MMSE with $G = 4$ achieves a significant performance gain over a 4×2 scheme-I system with $G = 2$, despite both having $N_r G = 4$. This is due to the decreased CCI following each (re-)transmission, as correctly decoded users are cancelled out resulting in a reduction in system loading. In a 4×1 system after $g = 3$ transmissions, primarily either none or one user is decoded in error; whereas in a 4×2 system after $g = 1$ transmission, the predominant trend includes either three or four users in error. Higher values of G allow more cancellation of correctly decoded users, thereby increasing the SINR and allowing us to apply linear MMSE for overloaded conditions. However, multiple retransmissions degrade throughput as will be seen in Figure 3.8.

Looking at Figures 3.3 and 3.4, it is observed that at high SNR, scheme-I for an overloaded 4×1 system using $G = 4$ achieves similar performance employing ML or MMSE. This is because at higher SNR values after sufficient transmissions, scheme-I results in a single user retransmission for both ML and MMSE MUD techniques.

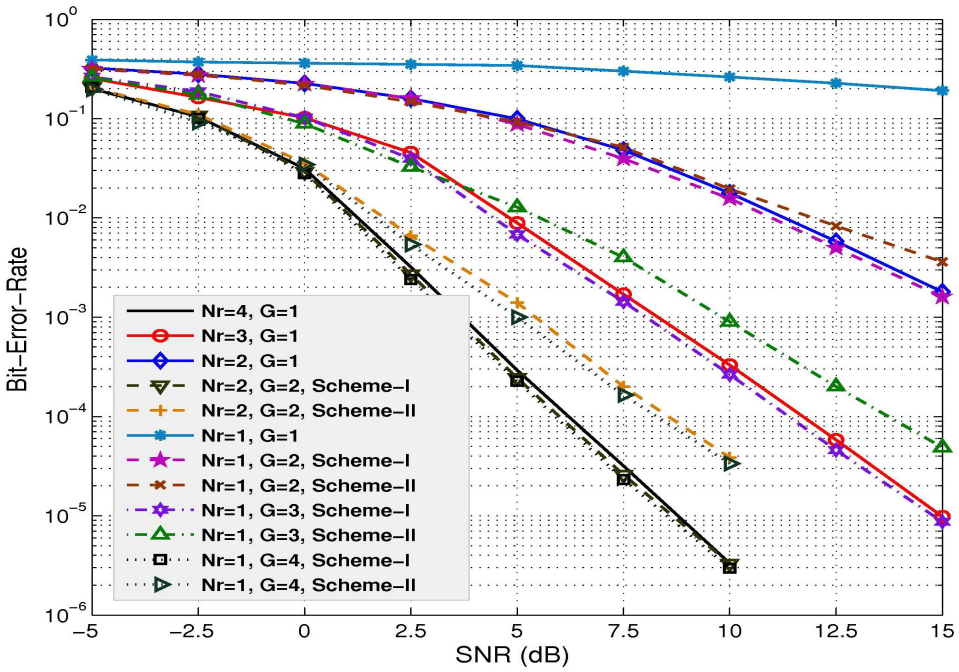


Figure 3.3: BER performance of proposed schemes with four users, $N_t = 4$, using ML MUD.

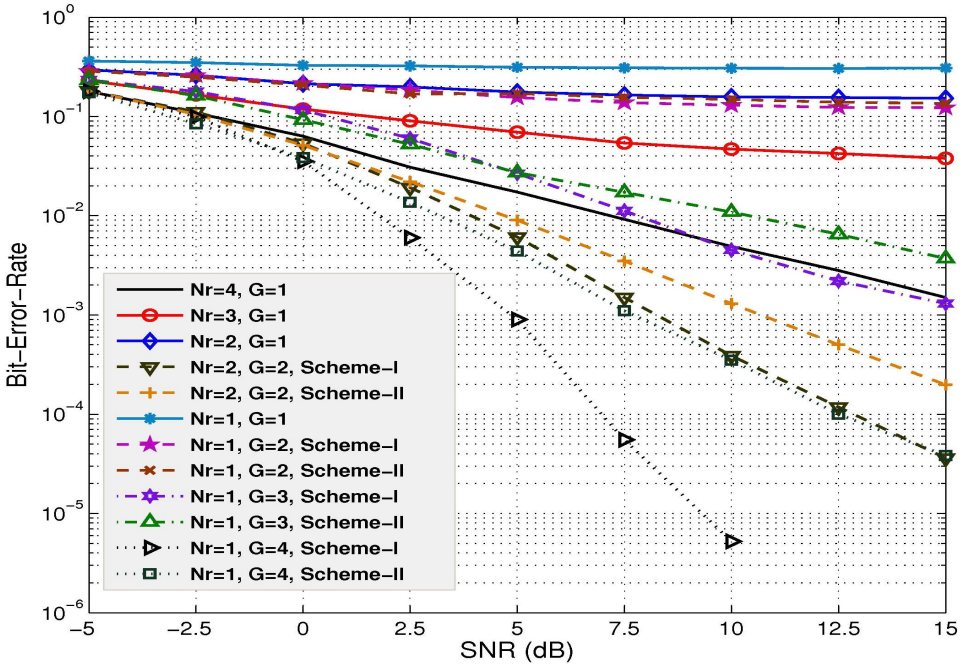


Figure 3.4: BER performance of proposed schemes with four users, $N_t = 4$, using linear MMSE MUD.

As expected, scheme-I achieves better performance than scheme-II for both ML and MMSE MUD techniques due to the decreasing CCI.

If a packet is still in error after some maximum number of transmissions, G, it is declared a dropped packet. The dropped packet-rate, $\mathcal{P}_{rate}$, which is a measure of the quality of the wireless system, is defined as the number of packets which are dropped divided by the total number of transmitted packets. Figures 3.5 and 3.6 compare the dropped packet-rate performance for both ML and linear MMSE MUD in the proposed two and four user schemes, respectively. Both figures show a similar performance pattern in terms of $\mathcal{P}_{rate}$ to what we have seen in terms of BER in Figure 3.2 for the two user case and in Figures 3.3 and 3.4 for four user case. As expected MMSE shows poorer $\mathcal{P}_{rate}$ performance than ML. On the other hand, MMSE has lower computational complexity than ML whose complexity grows exponentially with the number of transmit antennas and alphabet size.

The spectral efficiency, which is a measure of the rate of information transmission per unit bandwidth, can be combined with the dropped packet rate to define the so-called throughput spectral efficiency as [82]

$$\eta = \frac{\log_2 |C| R_c (1 - \mathcal{P}_{rate})}{N_{avg}} \quad \text{(b/s/Hz)}, \tag{3.10}$$

where N_{avg} is the average number of transmissions for each packet. The throughput spectral efficiency performance of soft output ML and linear MMSE MUD for the proposed two and four user HARQ system under different loading conditions is shown in Figure 3.7 and Figure 3.8, respectively. As expected, the overloaded schemes using $G > 1$ retransmissions, do not achieve the same throughput performance as the critically loaded system ($N_t = N_r$) with one transmission, $G = 1$. However, the schemes with $N_r < N_t$ do achieve significant performance gains using $G > 1$ rather than $G = 1$, mainly due to the fact the diversity gain achieved by the proposed schemes after HARQ retransmission improves the $\mathcal{P}_{rate}$ performance, thereby resulting in an increased throughput efficiency. Moreover, both figures clearly show that scheme-II achieves

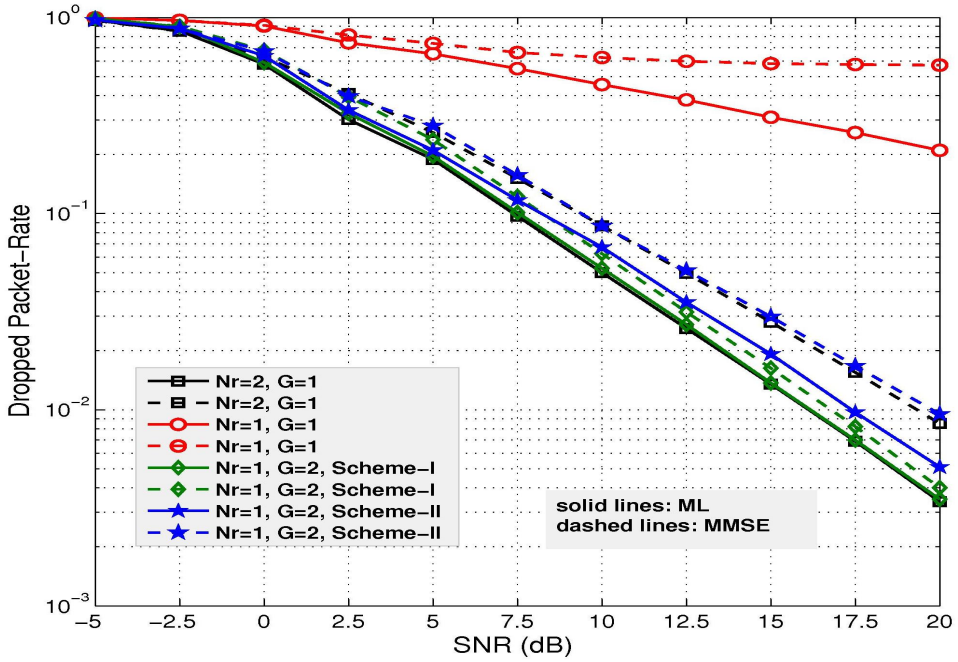


Figure 3.5: Dropped packet-rate performance of proposed two user schemes, $N_t = 2$, using ML (solid lines) and linear MMSE MUD (dashed lines).

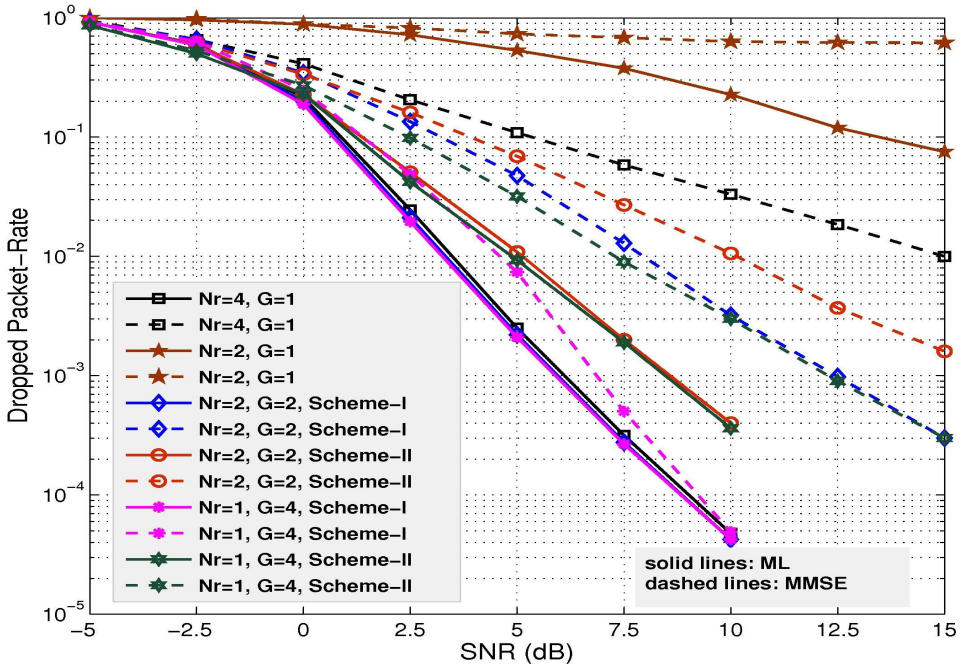


Figure 3.6: Dropped packet-rate performance of proposed four user schemes, $N_t = 4$, using ML (solid lines) and linear MMSE MUD (dashed lines).

marginal throughput gain over scheme-I at higher SNR values due to sending new information during retransmissions. However, the cost is increased CCI.

3.5 Chapter Conclusions

In this chapter, two detection schemes that combine HARQ with the use of virtual receive antennas for overloaded MIMO transmission systems are described. This combination results in a new systems approach that transforms an overloaded system into a critically or underloaded one. The proposed schemes allow us to apply linear MUD rather than employing high computational complexity MUD algorithms to handle overload. Simulations demonstrate that the new schemes can result in significant gains in terms of BER and dropped packet performance. Moreover, these schemes, especially Scheme-II, can increase the throughput spectral efficiency of a system under overloaded conditions.

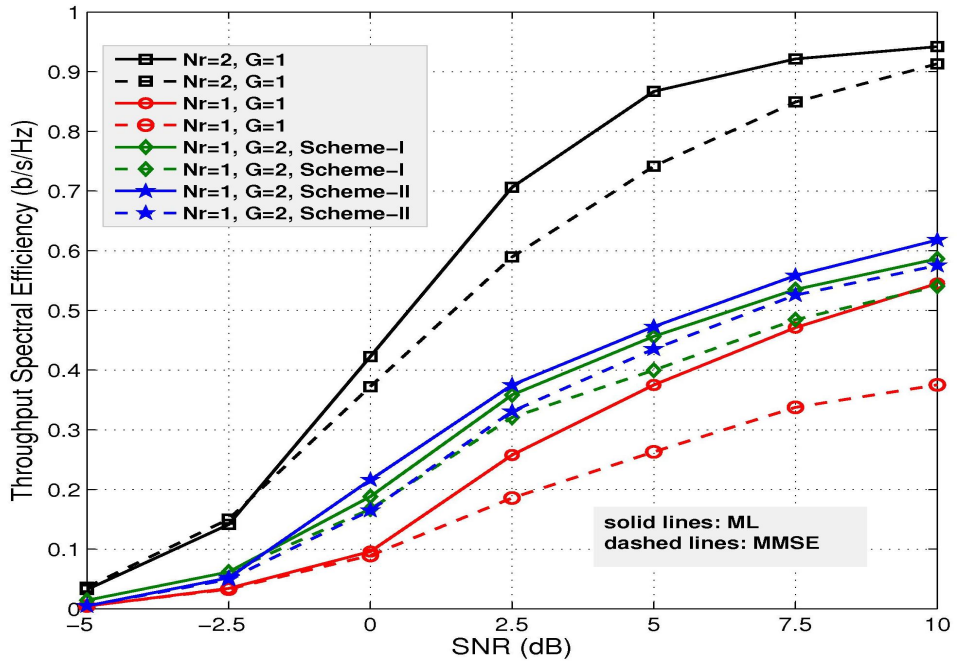


Figure 3.7: Throughput spectral efficiency of proposed two user schemes, $N_t = 2$, using ML (solid lines) and linear MMSE MUD (dashed lines).

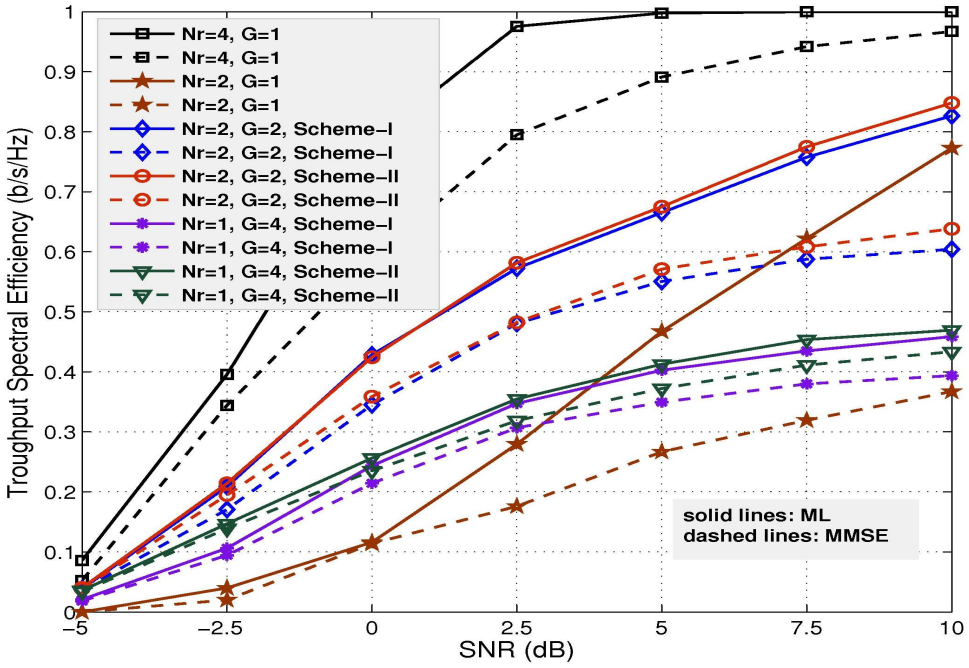


Figure 3.8: Throughput spectral efficiency of proposed four user schemes, $N_t = 4$, using ML (solid lines) and linear MMSE MUD (dashed lines).

Chapter 4

Cooperative DF Relaying Systems

4.1 Introduction

Hybrid automatic repeat request (HARQ) is a common technique that has been adopted in emerging wireless standards such as worldwide interoperability for microwave access (WiMAX) [83] and long term evolution (LTE) [84] to provide reliable communication over noisy wireless channels. In [34], it was suggested that integrating HARQ techniques into cooperative relaying systems may lead to significant improvements in terms of reliable decoding at the destination, throughput efficiency and energy consumption compared to conventional point-to-point HARQ systems. Since then, several researchers have investigated the advantages of HARQ protocols in cooperative relaying networks [33], [35]-[41].

However, all these studies assumed that there is always a direct link available between source and destination. Therefore, when there is an erroneous packet at the destination, a retransmission can come from either the source or the relay(s) or both. Relatively few works [43]-[47] on multi-hop decode-and-forward (DF) relaying systems with HARQ have focused on the case of no direct link between the source and destination due to large physical separation or path loss. In the absence of a direct link,

relay(s) adopting a half-duplex DF relaying protocol will request retransmission whenever they receive an erroneous packet. An important limitation of [43]-[47] and similar studies is that they have considered only a single source that is assisted by a single or multiple relays. The extent to which the above results may be generalized to multiple source scenarios remains unclear.

In this chapter, we focus on the more challenging case of multi-source and multi-relay configuration for two-hop cooperative DF relaying network with no direct link. The existence of multiple sources and relays raises questions on which source nodes need to retransmit to whom, when and how. To address this, a multi process HARQ retransmission scheme is developed by allowing simultaneous retransmissions from source node(s) of packets detected in error. This results in improved throughput compared to traditional orthogonal (time division) retransmissions. To avoid unnecessary retransmissions, upon completion of each (re-)transmission round from the source nodes, relays are allowed to exchange their decoding outcomes. Moreover, a novel forwarding strategy with minimum overhead is proposed with the objective of further improving the throughput. The proposed forwarding scheme does not require any centralization or participation either from the sources or the destination.

The performance analysis in terms of outage probability and throughput of the proposed retransmission and relay forwarding schemes is presented in Chapter 5. This chapter is organized as follows. In Section 4.2, the system model for a two-hop DF relaying network with no direct link is described. Section 4.3, introduces the proposed retransmission scheme during the broadcast phase. In addition, a forwarding strategy to support simultaneous transmissions from relays is presented. Finally, conclusions are drawn in Section 4.4.

4.2 System Model

Consider an uplink system consisting of two co-channel source nodes, $S_u, u = 1, 2$, communicating with a destination, D, through two relays, $R_l, l = 1, 2$ as shown in

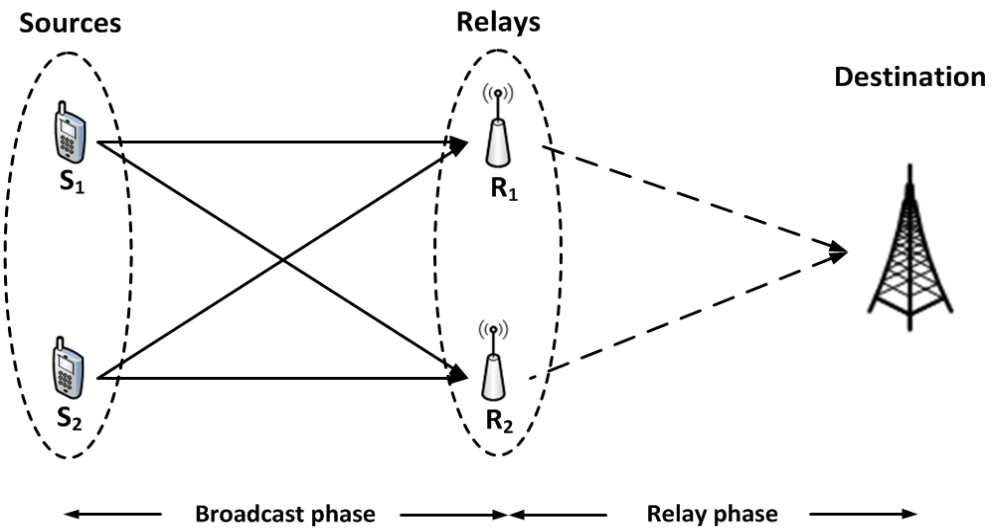


Figure 4.1: Two-hop relaying network with two source and two relay nodes.

Figure 4.1. Each source and relay node is equipped with a single antenna. The destination, typically assumed to be a base station, is equipped with N_d receive antennas. Transmission of data packets from source nodes to the destination operates in two stages, a broadcast phase (solid lines) and a relay phase (dashed lines). Any direct link between sources and destination is ignored due to the (assumed) large physical separation between them. Each relay adopts a half-duplex DF relaying protocol [35]. As in [85], all links are assumed to be short-term quasi-static fading, meaning that the channel coefficients remain constant during a (re-)transmission round of length L, but vary independently between rounds. The channel between two nodes a and b (i.e. uth source - lth relay or lth relay - destination) is assumed to exhibit frequency-flat Rayleigh fading. The channel coefficients, $h_{a,b}(g)$, between two nodes a and b at the gth (re-)transmission round are modeled as independent identically distributed (i.i.d) complex Gaussian random variables with zero-mean and variance $\sigma^2_{a,b}$. Finally, it is assumed that perfect channel state information (CSI) is available for the source-relay links at the relays, and for the relay-destination links at the destination.

4.3 (Re-)transmission and Forwarding Schemes

This section presents the proposed retransmission and forwarding schemes during the broadcast phase and relay phase, respectively.

4.3.1 Broadcast Phase

During the first transmission round, each source node transmits its pth data packet to both relays in time division multiple access (TDMA) fashion, with $p \in \{1, 2, \ldots\}$. The resulting received signal $\mathbf{y}_l(1)$ at relay R_l can be written as

$$\mathbf{y}_l(1) = h_{u,l}(1)\ \mathbf{x}_u + \mathbf{v}_l(1), \tag{4.1}$$

where $h_{u,l}(1)$ is the channel coefficient between the uth source node and the lth relay, $\mathbf{x}_u \in \mathbb{C}^{1\times L}$ is the transmitted symbol sequence from source S_u and $\mathbf{v}_l(1) \in \mathbb{C}^{1\times L}$ is a complex additive white Gaussian noise (AWGN) vector with variance $\sigma^2_{v,l}$ at R_l. The instantaneous received signal-to-noise ratio (SNR) at relay node R_l during the first transmission is given by

$$\gamma_{u,l}(1) = \frac{E_S|h_{u,l}(1)|^2}{N_0}, \tag{4.2}$$

where E_S is the average symbol energy transmitted from each source. Let the average received SNR at relay R_l be defined as $\overline{\gamma}_{u,l} = E\{\gamma_{u,l}(1)\}$.

After completion of the first transmission from both source nodes, the relays R_1 and R_2 take turns to broadcast two-bit ACK(s) and/or NACK(s) message to the sources and to each other. Hence, a relay is aware not only of its own decoding outcomes, but also has information about the success or failure of signals received at the other relay. It is assumed that S_u will retransmit a data packet only if it receives a NACK from both R_1 and R_2. Let $\mathcal{U}_l$ denotes the set of source nodes that are successfully decoded at relay R_l, and let $\mathcal{S}$ and $\mathcal{V}$ denotes the set of source nodes that are successfully and erroneously decoded at both relays, respectively.

A few key assumptions are made:

1. All feedback channels are error free.

2. Both source nodes are assumed to be at the same distance from the lth relay, and both relays are placed at the mid-point between the source nodes and destination. We then define $\overline{\gamma}_{u,l} = \overline{\gamma}_{l,D} = \rho$, where $\overline{\gamma}_{l,D}$ is the average received SNR at the destination.

3. The distance between the relays is much shorter than the distance between any two source-relay nodes. Therefore, the ACK/NACK channel between the relays will experience lower propagation delay compared to that due to the feedback channel from relays to sources.

4. Only one retransmission[1] is allowed for each source node's data packet. Depending upon the instantaneous link quality, the single retransmission may be needed either during the broadcast or relay phase. Let G denote the maximum number of transmissions for each source node's data packet in a two-hop DF relaying system, which is equal to the sum of transmission rounds in each transmission phase. Therefore, $G = 3$, for each source node's data packet.

Two different retransmission schemes to describe the retransmission process in the broadcast phase are outlined as follows.

Orthogonal (time division) retransmission

When the data packets of both source nodes are detected in error at both relays, S_1 and S_2 will retransmit their erroneously received data packets during the next two time slots. Relay R_l will combine the retransmitted information and the previous received information from each source node using maximal ratio combining (MRC) [86]. The resulting MRC combined vector $\mathbf{y}_{l,2}$ at relay R_l after two transmissions of the same

[1]The extension to more than one retransmissions for each source node is straightforward. However, allowing multiple retransmissions can degrade throughput.

packet from source S_u is given by

$$\mathbf{y}_{l,2} = \sum_{g=1}^{2} \alpha(g)\mathbf{y}_l(g), \tag{4.3}$$

where $\alpha(g) = h^*_{u,l}(g)/N_0$ is the combining coefficient. If any of the retransmitted packets are still in error at both relays then the erroneous packet is dropped. The relay(s) will forward only the data packet of a successfully decoded source node during the relay phase.

When only one source node's data packet is successfully decoded at one or more relays, only the unsuccessful source will retransmit the same data packet while the other source remains silent.

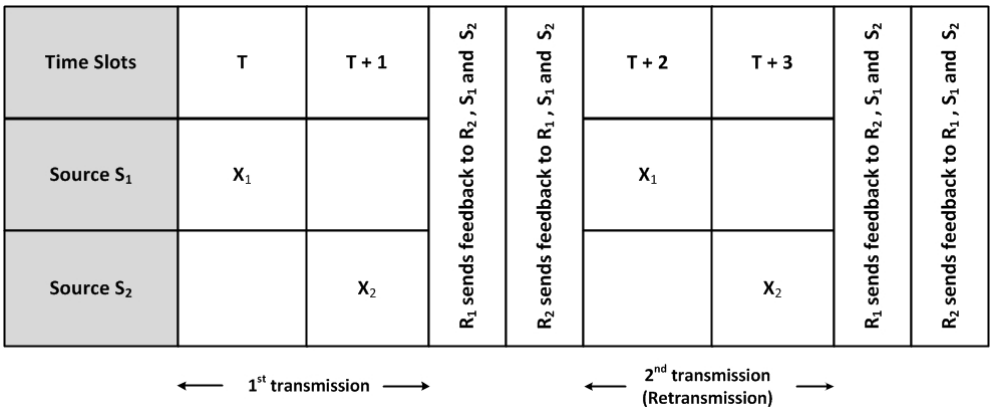
Proposed retransmission scheme

Here, we allow simultaneous retransmissions from the source nodes to the relays. Retransmitted information and previously received information from both source nodes in two different time slots will be stacked and then treated as if there were three virtual receive antennas [21, 23] at each relay. Let N_t denote the total number of transmit antennas across all source nodes, and let $N_l = 3$, denote the number of virtual receive antennas at each relay after two transmissions. The resulting stacked matrix at relay R_l can be written as

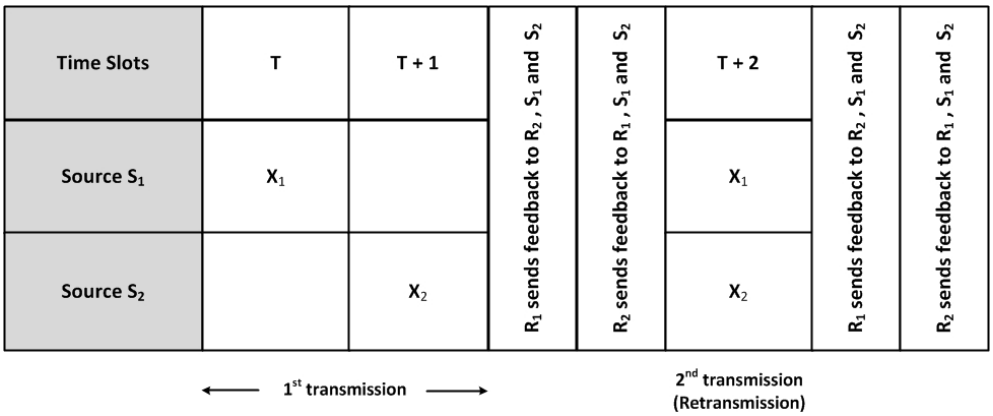
$$\mathbf{R}_{l,2} = \begin{bmatrix} h_{1,l}(1) & 0 \\ 0 & h_{2,l}(2) \\ h_{1,l}(2) & h_{2,l}(2) \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{v}_l(1) \\ \mathbf{v}_l(1) \\ \mathbf{v}_l(2) \end{bmatrix} = \bar{\mathbf{H}}_l(2)\mathbf{X} + \bar{\mathbf{V}}_l(2), \tag{4.4}$$

where $\bar{\mathbf{H}}_l(2) \in \mathbb{C}^{N_l \times N_t}$ is the equivalent channel matrix, $\mathbf{X} \in \mathbb{C}^{N_t \times L}$ is the overall transmitted sequence from both source nodes and $\bar{\mathbf{V}}_l(2) \in \mathbb{C}^{N_l \times L}$ is the equivalent noise matrix at relay R_l. Due to the computational limitations of the relay nodes in our system model, a relay performs only linear multi-user detection (MUD) on the

resulting stacked matrix. If any packets are still in error at both R_1 and R_2 then the erroneous packets are dropped.



(a) Orthogonal (time division) retransmission.



(b) Proposed retransmission scheme.

Figure 4.2: Retransmission schemes during the broadcast phase. Empty blocks denote that a source node remains idle during that time.

Orthogonal (time division) retransmissions from both source nodes would require two time slots to complete the HARQ retransmission round. Relaxing the orthogonality requirements and allowing source nodes to retransmit simultaneously requires only one additional time slot to complete the HARQ retransmissions. An example of the orthogonal retransmission approach and the proposed retransmission scheme is shown in Figure. 4.2. During time slots T and $T+1$, S_1 and S_2 transmit their data packets while R_1 and R_2 listen. During the feedback period, we assume that first R_1 broadcasts

two-bit NACKs to S_1, S_2 and R_2, then R_2 broadcasts two-bit NACKs to S_1, S_2 and R_1. For the orthogonal retransmissions, S_1 will retransmit during time slot $T+2$ and S_2 in time slot $T+3$ as shown in Figure 4.2(a). On the other hand, for the proposed retransmission scheme, S_1 and S_2 will simultaneously retransmit during time slot $T+2$. At the end of HARQ retransmission, R_1 and R_2 take turns to broadcast their decoding outcomes (ACK/NACK). Figure 4.3 illustrates the flow chart of the (re-)transmission process during the broadcast phase.

The exchange of decoding outcomes among relay nodes upon completion of each (re-)transmission round from the source nodes allows us to avoid unnecessary retransmissions, thereby improving the average retransmission rate. To get further insight, we look at some simulation results, obtained via averaging over 15000 channel realizations. Consider quasi-static Rayleigh frequency-flat fading channel and 4-quadrature amplitude modulation (QAM) modulation.

Figure 4.4 illustrates the average retransmission rate after completion of the first transmission round from both source nodes. As discussed earlier, the first transmission from both source nodes is performed over orthogonal channels (i.e. TDMA). First, we evaluate the first transmission decoding outcomes at an individual relay, let say R_1. From Figure 4.4(a), it is observed that at lower SNR (e.g. 0dB), 56% of the time both sources will retransmit their data packets and 41% of the time only one source node retransmits its data packet. As the SNR increases, the retransmission rate of the two source nodes as well as single source node decreases, while there is an increase in the rate of no source node retransmits its data packet. For example at 15dB, 89% of the time R_1 does not request retransmission from any source node.

We now evaluate the first transmission decoding outcomes over both relays. From Figure 4.4(b), we can see that at low SNR, the retransmission rate of the two source nodes decreases from 56% to 33%, while there is an increase from 41% to 60% in the retransmission rate of a single source compared to outcomes only at R_1. At high SNR, the retransmission rate of two and single source node decreases and the rate of no

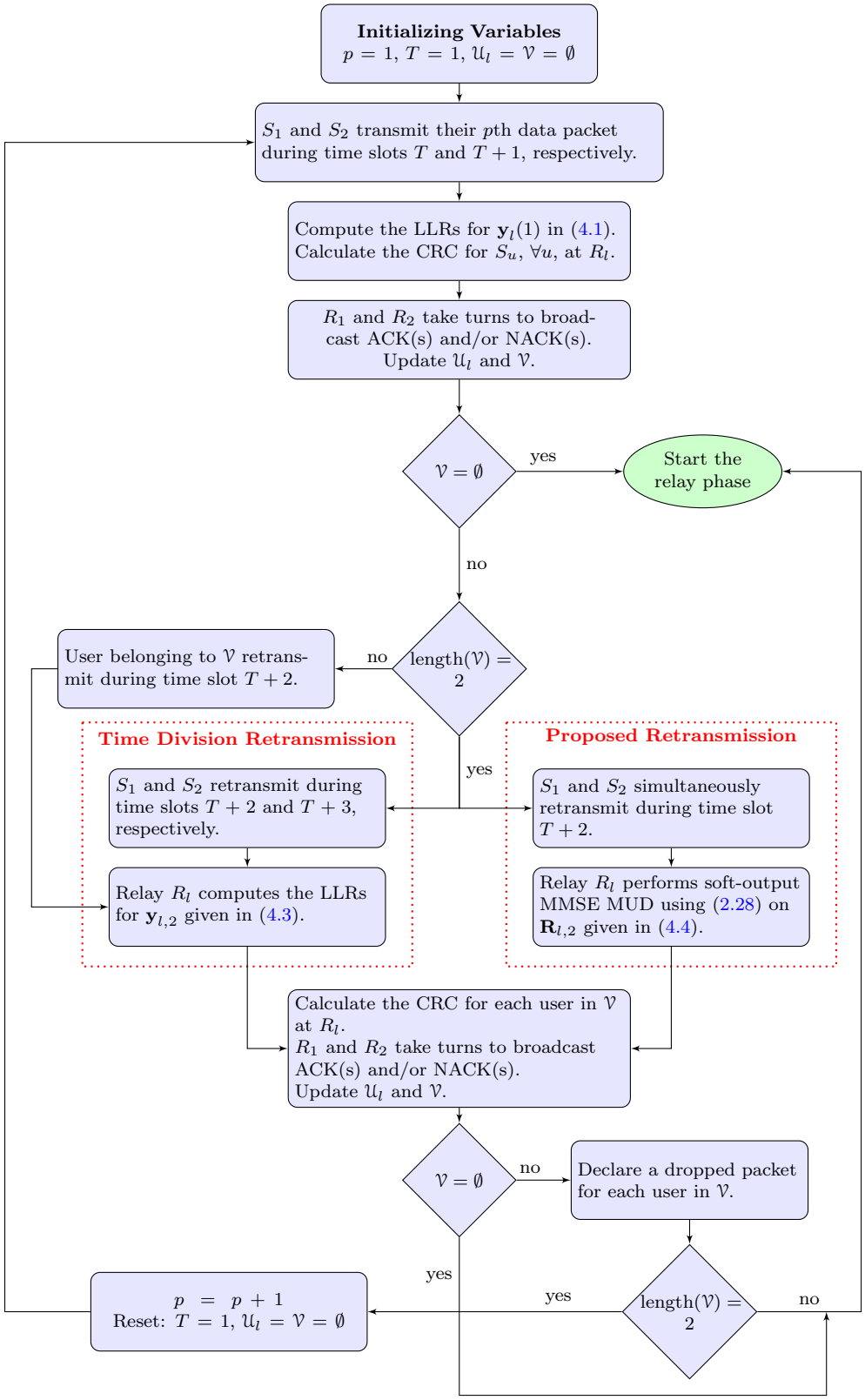
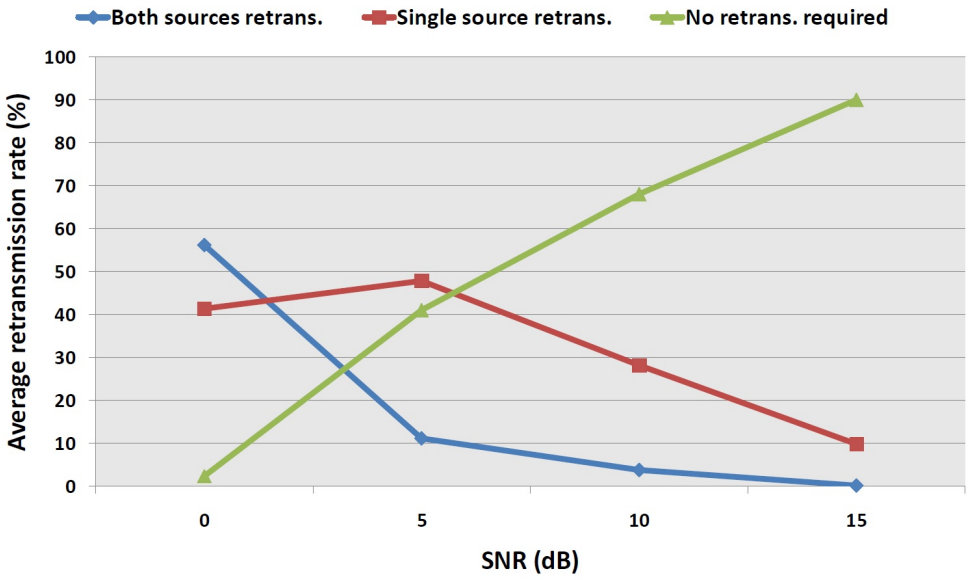
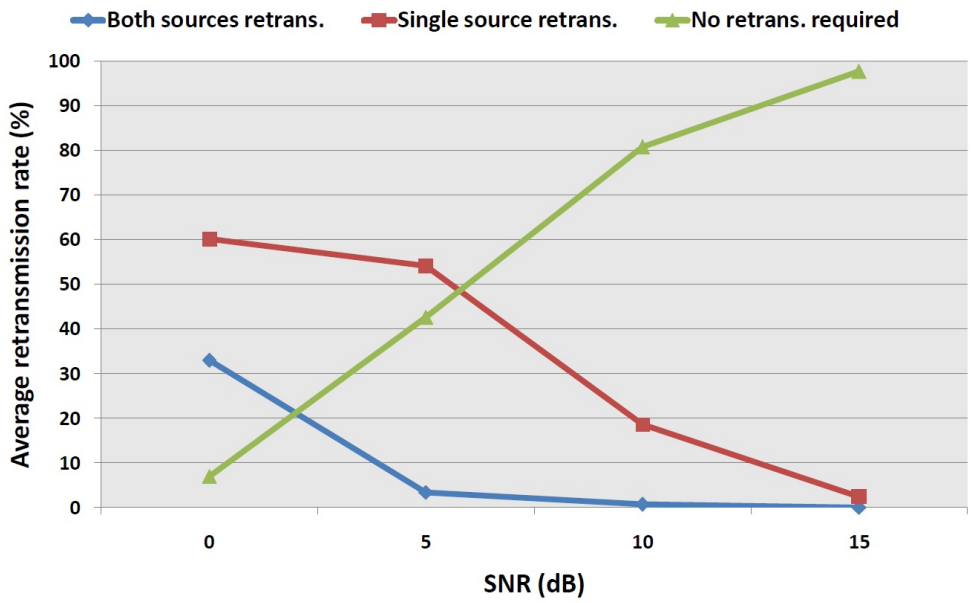


Figure 4.3: Flow chart of HARQ retransmission schemes during the broadcast phase.

retransmissions required increases.



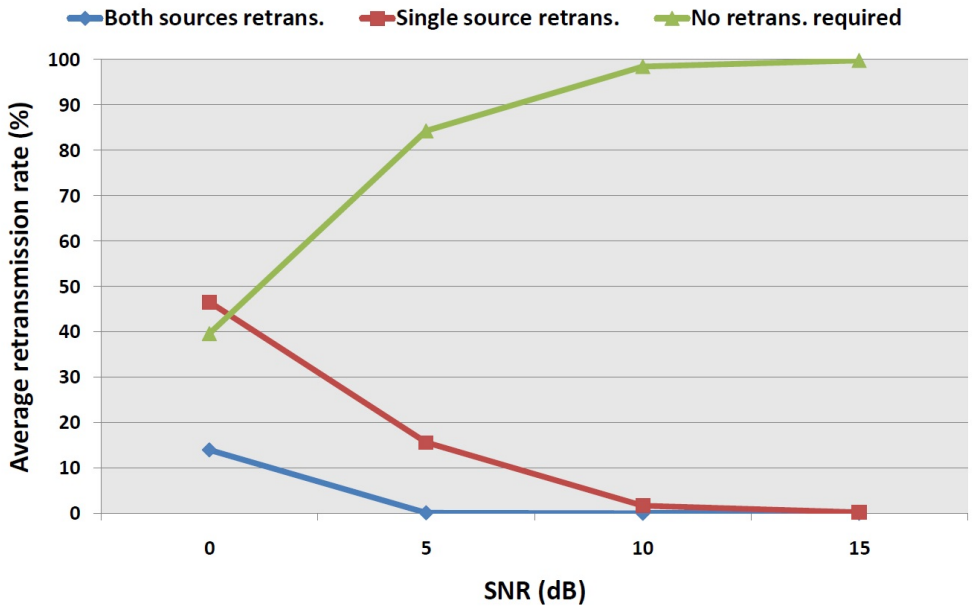
(a) Orthogonal (time division) first transmission decoding outcomes at R_1.



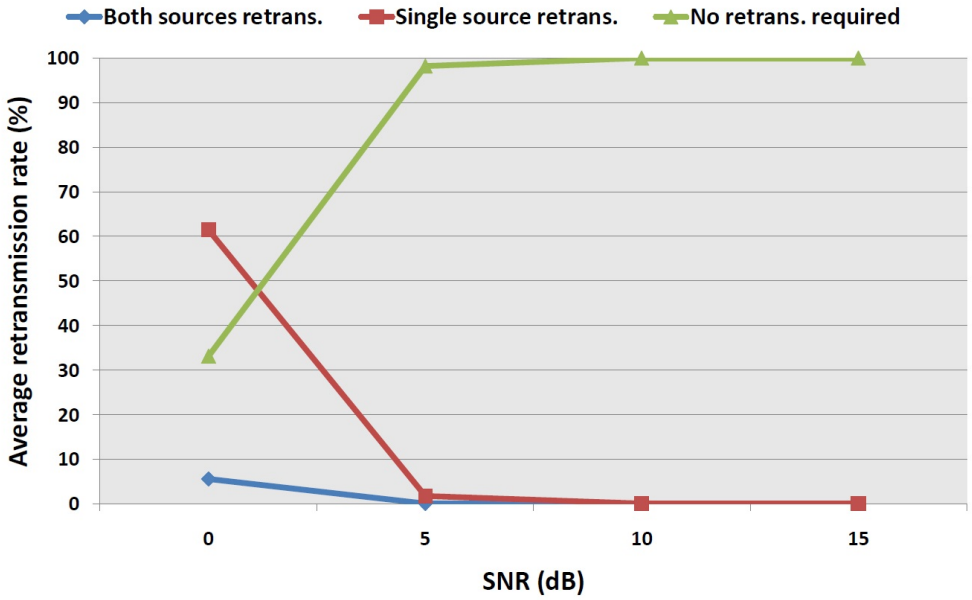
(b) Orthogonal (time division) first transmission decoding outcomes over both relays.

Figure 4.4: Average retransmission rates for orthogonal (time division) first transmissions during the broadcast phase.

Figures 4.5 and 4.6 present the average retransmission rate after two transmissions for the orthogonal (time division) retransmission and the simultaneous (proposed) re-



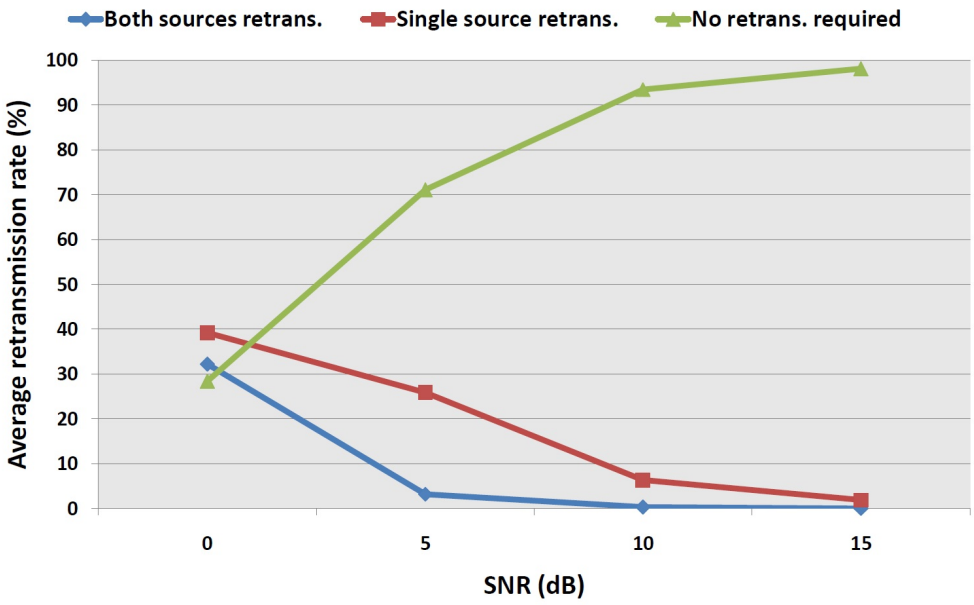
(a) Orthogonal (time division) retransmission decoding outcomes at R_1.



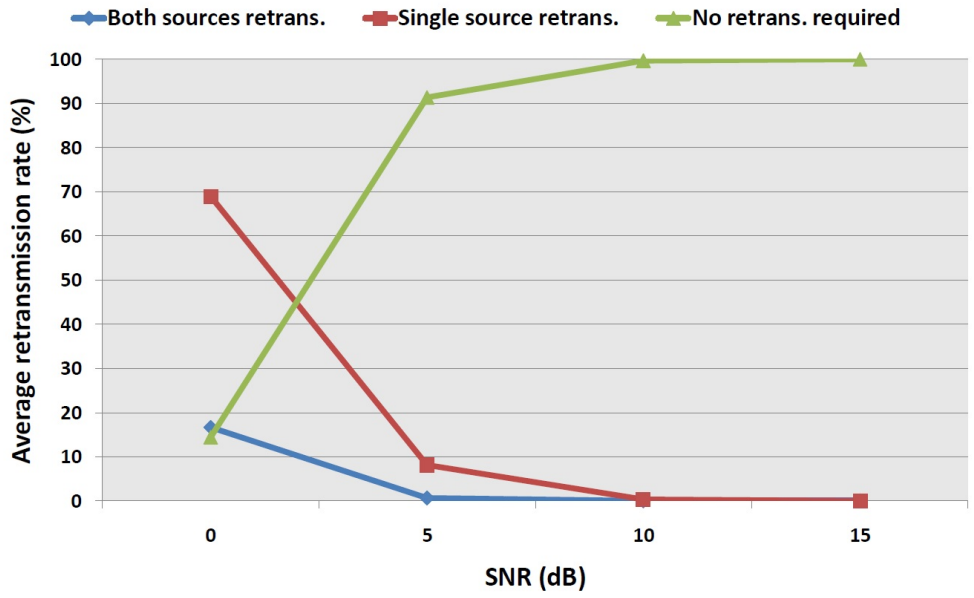
(b) Orthogonal (time division) retransmission decoding outcomes over both relays.

Figure 4.5: Average retransmission rates for orthogonal (time division) retransmission during the broadcast phase.

transmission schemes, respectively. From both figures, it is observed that as the number of HARQ rounds increases, the retransmission rate for two source nodes and single



(a) Proposed retransmission decoding outcomes at R_1.



(b) Proposed retransmission decoding outcomes over both relays.

Figure 4.6: Average retransmission rates for proposed retransmission during the broadcast phase.

source node decreases, while the rate of no retransmissions required increases. Moreover, comparing the two figures, we note that at high SNR, the average retransmission rate for both retransmission schemes is almost similar.

4.3.2 Relay Phase

After the gth transmission of the broadcast phase, where $g = 1, 2, \ldots, G-1$, both relays have information about each other's decoding outcomes. Based on this assumption, a novel forwarding strategy is proposed in order to maximize the number of simultaneous transmissions during the relay phase, with the objective of improving the throughput.

The basic idea behind the proposed strategy is to assign each relay one priority source node. Without loss of generality, S_1 is assigned as the priority source node of R_1 and S_2 as the priority source node of R_2. Relay R_l is declared an active relay if $\mathcal{U}_l \neq \emptyset$. Let w be the number of active relays and the set of active relays be denoted by $\mathcal{N} \subset \mathcal{F}$, where $\mathcal{F}$ is the set of all relays and $|\mathcal{N}| = w$. We look at two cases to describe the forwarding strategies for the relay phase.

Case-I (Both relays participate in the relay phase)

During the first transmission of the relay phase, both relays will simultaneously transmit the data packet of their respective priority source node. If R_1's and R_2's decoding sets contain each others priority source node instead of their own (e.g. $\mathcal{U}_1 = \{S_2\}$ and $\mathcal{U}_2 = \{S_1\}$), then the relays will independently decide to swap their priority source nodes and simultaneously transmit the data packets to the destination. Switching of priority source nodes is performed at the relays without the participation of either the sources or the destination, based on information obtained from the relay ACK(s)/NACK(s).

The received signal $\mathbf{Y}_D(1)$ at the destination during the first transmission is given by

$$\mathbf{Y}_D(1) = \sum_{l=1}^{w} \mathbf{h}_l(1)\ \mathbf{x}_l + \mathbf{V}_D(1), \tag{4.5}$$

where $\mathbf{h}_l(1)$ is the lth column of the channel matrix $\mathbf{H}_D(1)$ of size $N_d \times M$ with entries, $h_{l,d}(1)$, corresponding to the channel coefficient between the l^{th} relay and the d^{th} receive antenna, M is the total number of antennas across all relays, $\mathbf{x}_l \in \mathbb{C}^{1\times L}$ is the transmitted symbol sequence from relay R_l and $\mathbf{V}_D(1) \in \mathbb{C}^{N_d \times L}$ is a complex additive white Gaussian noise (AWGN) matrix with entries having variance $\sigma^2_{v,D}$. The instan-

taneous and average received SNRs at the destination during the first transmission of the relay phase are given by $\gamma_{l,D}(1) = \frac{E_R \|\mathbf{h}_l(1)\|_F^2}{N_0}$ and $\overline{\gamma}_{l,D} = E\{\gamma_{l,D}(1)\}$, where $\|\cdot\|_F^2$ is the Frobenius norm, $\|\mathbf{h}_l(1)\|_F^2 = \sum_{d=1}^{N_d} |h_{l,d}(1)|^2$, and E_R is the average symbol energy transmitted from each relay.

The destination employs either optimal ML detection or a suboptimal linear MUD algorithm to reduce the co-channel interference (CCI) due to simultaneous transmissions from the relays. After the CRC check at the destination, ACK(s) and/or NACK(s) will be broadcast by the destination to the relays. Since, only one retransmission is allowed for each source node's data packet, retransmission of an erroneous data packet during the relay phase will only be possible if it is received correctly at a relay during its initial transmission round. Based on the decoding outcomes at the destination, three different cases are possible:

- If the destination broadcasts two ACK messages, both relays will stop the relay phase and request the source nodes to transmit their next data packet.
- If the destination broadcasts two NACK messages, each relay will simultaneously retransmit its corresponding source nodes data packet. Retransmitted information will be treated as creating an additional N_d virtual receive antennas. Once the relay phase is complete, both relays will request the source nodes to transmit their next data packet.
- If the destination broadcasts an ACK and a NACK, the relay whose transmitted data packet was successfully decoded at the destination will remain silent while the relay whose transmitted data packet was erroneous will retransmit in the next time slot.

Case-II (Only one relay participates in the relay phase)

Case-II is considered if any of the following conditions are met:

- All relay decoding sets contain only S_1 or S_2 at the end of the broadcast phase,

e.g. $\mathcal{U}_1 = \{S_1\}$ and $\mathcal{U}_2 = \{S_1\}$. In that case, the relay R_1 with priority source node S_1 will forward the information to the destination, while the other relay R_2 will remain silent. At the end of the first transmission, the destination will feedback an ACK/NACK message to the relay. On receiving a NACK, the same relay will retransmit the same data packet in the next time slot.

- $w = 1$, i.e. only one active relay is available to forward a source node(s) data packet to the destination over orthogonal channels. Once the relay phase is complete, the relay will request the source nodes to transmit their next data packet.

Table 4.1 further elaborates the relay forwarding strategies for the two-hop DF relaying system presented in Section 4.2. "✓"and "⊘"indicates whether the source node S_u, $u = 1, 2$, belongs or does not belong to set $\mathcal{U}_l$, $l = 1, 2$, respectively. "$R_l \longrightarrow S_u$"represents that relay R_l forwards the S_u's data packet.

Scenario No.	$\mathcal{U}_1$		$\mathcal{U}_2$		**Relay Forwarding Strategy**	
	S_1	S_2	S_1	S_2		
1	✓	✓	✓	✓	Case-I	$R_1 \longrightarrow S_1$ and $R_2 \longrightarrow S_2$
2	✓	✓	⊘	✓	Case-I	$R_1 \longrightarrow S_1$ and $R_2 \longrightarrow S_2$
3	✓	⊘	✓	✓	Case-I	$R_1 \longrightarrow S_1$ and $R_2 \longrightarrow S_2$
4	✓	⊘	⊘	✓	Case-I	$R_1 \longrightarrow S_1$ and $R_2 \longrightarrow S_2$
5	✓	✓	✓	⊘	Case-I	$R_1 \longrightarrow S_2$ and $R_2 \longrightarrow S_1$
6	⊘	✓	✓	✓	Case-I	$R_1 \longrightarrow S_2$ and $R_2 \longrightarrow S_1$
7	⊘	✓	✓	⊘	Case-I	$R_1 \longrightarrow S_2$ and $R_2 \longrightarrow S_1$
8	✓	✓	⊘	⊘	Case-II	$R_1 \longrightarrow S_1, S_2$ and R_2 remains silent
9	✓	⊘	⊘	⊘	Case-II	$R_1 \longrightarrow S_1$ and R_2 remains silent
10	⊘	✓	⊘	⊘	Case-II	$R_1 \longrightarrow S_2$ and R_2 remains silent
11	⊘	⊘	✓	✓	Case-II	$R_2 \longrightarrow S_1, S_2$ and R_1 remains silent
12	⊘	⊘	✓	⊘	Case-II	$R_2 \longrightarrow S_1$ and R_1 remains silent
13	⊘	⊘	⊘	✓	Case-II	$R_1 \longrightarrow S_2$ and R_1 remains silent
14	✓	⊘	✓	⊘	Case-II	$R_1 \longrightarrow S_1$ and R_2 remains silent
15	⊘	✓	⊘	✓	Case-II	$R_2 \longrightarrow S_2$ and R_1 remains silent

Table 4.1: Proposed forwarding strategies during the relay phase.

4.4 Chapter Conclusions

In this chapter, multiple HARQ processes were studied during both the broadcast and the relay phases for multi-source multi-relay DF relaying networks with no direct link, a topic that has not been investigated in the literature. Dealing with multiple HARQ processes at each relay, a retransmission scheme was proposed that utilized virtual antennas to achieve increased receive diversity and improved throughput compared to traditional orthogonal (time division) retransmissions. Exchange of decoding outcomes (ACK(s)/NACK(s)) among relay nodes was allowed which resulted in an improved retransmission rate as demonstrated by the simulation results. Based on this assumption a novel forwarding strategy to maximize the throughput was also proposed in a very simple way, which highlights the advantages of the proposed forwarding strategy. The numerical results in terms of outage probability and throughput for the proposed schemes are presented in Chapter 5.

Chapter 5

Performance Analysis of Cooperative DF Relaying Systems

This chapter presents a theoretical analysis of the two-hop DF relaying system introduced in Chapter 4. First, the outage probabilities of the proposed retransmission and forwarding schemes are evaluated in Section 5.1. The outage probability is defined [26, 72] as the probability that the mutual information between two nodes falls below a specified target transmission rate[1] $\mathcal{R}$ (bits/s/Hz). It can then be expressed as [72]

$$P^{\text{out}} = \Pr\Big(I_{a,b} < \mathcal{R}\Big), \tag{5.1}$$

where $I_{a,b}$ is the mutual information of a channel with input a and output b. Using these, expressions for the throughput of the proposed retransmission and forwarding schemes are derived in Section 5.2. In Section 5.3, the numerical results are presented and finally the conclusions are drawn in Section 5.4.

[1] For simplicity, it is assumed that all transmitting nodes (source/relay) transmit at the same rate.

5.1 Outage Probability Analysis

In this section, we first analyze the outage probability for each transmission phase. We then find the end-to-end (e2e) outage probability (the probability that at least one transmission phase, either broadcast or relay, is in outage) for various (re-)transmission modes. The e2e outage probability is an important measure of the reliability of a multi-hop DF relaying system in the absence of a direct link between source node(s) and destination(s).

5.1.1 Broadcast Phase Outage Probability

We begin by evaluating the outage probability after the first transmission round of the broadcast phase. As discussed in Section 4.3.1, the first transmission from the two source nodes is carried out over two successive time slots. The mutual information between the uth source and lth relay in the first transmission is given by [35]

$$I_{u,l}^{\text{tdma}}(1) = \log_2\left(1 + |h_{u,l}(1)|^2 \frac{E_S}{N_0}\right), \tag{5.2}$$

where $u = 1, 2$, $l = 1, 2$ and N_0 denotes the one-sided noise spectral density. The outage probability after the first transmission is then [35]

$$P_{u,l}^{\text{out,tdma}}(1) = \Pr\Big(I_{u,l}^{\text{tdma}}(1) < \mathcal{R}\Big). \tag{5.3}$$

From (5.2) and (5.3), we obtain

$$\begin{aligned} P_{u,l}^{\text{out,tdma}}(1) &= \Pr\left(\log_2\left(1 + |h_{u,l}(1)|^2 \frac{E_S}{N_0}\right) < \mathcal{R}\right) \\ &= \Pr\left(|h_{u,l}(1)|^2 \frac{E_S}{N_0} < \left(2^{\mathcal{R}} - 1\right)\right). \end{aligned} \tag{5.4}$$

If $\gamma_{u,l}(1) = \frac{E_S |h_{u,l}(1)|^2}{N_0}$ is the instantaneous received SNR at relay R_l during the first transmission, we can rewrite the outage probability given in (5.4) as [87]

$$\begin{aligned} P_{u,l}^{\text{out,tdma}}(1) &= \Pr\Big(\gamma_{u,l}(1) < \big(2^{\mathcal{R}} - 1\big)\Big) \\ &= \int_0^{2^{\mathcal{R}}-1} f(\gamma_{u,l}(1)) d\gamma_{u,l}(1), \end{aligned} \tag{5.5}$$

where $f(\gamma_{u,l}(1))$ denotes the probability density function (pdf) of $\gamma_{u,l}(1)$. For a Rayleigh fading channel, the instantaneous received SNR has an exponential distribution and therefore the outage probability can be expressed as [87]

$$P_{u,l}^{\text{out,tdma}}(1) = \int_0^{2^{\mathcal{R}}-1} \frac{1}{\rho} \exp\left(-\frac{\gamma_{u,l}(1)}{\rho}\right) d\gamma_{u,l}(1), \tag{5.6}$$

where $\rho = \frac{E_S}{N_0}$ is average received SNR at R_l. Applying the integration in (5.6) and after some manipulations, the $P_{u,l}^{\text{out,tdma}}(1)$ can be expressed as [87]

$$P_{u,l}^{\text{out,tdma}}(1) = 1 - \exp\left(-\frac{\big(2^{\mathcal{R}} - 1\big)}{\rho}\right). \tag{5.7}$$

Based on the outage events, relay R_l will decide whether to send an ACK/NACK after the first transmission. Since both relays exchange decoding outcomes (ACK/NACK), a retransmission from the uth source occurs only when an outage occurs at both relays, which means both $\Big(I_{u,1}^{\text{tdma}}(1) < \mathcal{R}\Big)$ and $\Big(I_{u,2}^{\text{tdma}}(1) < \mathcal{R}\Big)$.

Consider the worst case scenario, where the mutual information for every source-relay link during the initial transmission falls below the target rate. Both source nodes must then retransmit the data packets in their allocated time slots. In the following, the outage probability after two transmissions (including the retransmissions from both source nodes) is evaluated for both the orthogonal and proposed retransmission schemes.

a) *Orthogonal (time division) retransmission*: The relay R_l performs maximum ratio combining (MRC) on the received packets from each source node after two trans-

missions. The mutual information between the uth source and lth relay in the second transmission (i.e. $g = 2$) is given by [86]

$$I_{u,l}^{\text{tdma}}(2) = \log_2\left(1 + \sum_{m=1}^{2} |h_{u,l}(m)|^2 \frac{E_S}{N_0}\right). \tag{5.8}$$

The resulting outage probability is given by

$$\begin{aligned} P_{u,l}^{\text{out,tdma}}(2) &= \Pr\left(I_{u,l}^{\text{tdma}}(2) < \Re\right) \\ &= \Pr\left(\sum_{m=1}^{2} |h_{u,l}(m)|^2 \frac{E_S}{N_0} < \left(2^{\Re} - 1\right)\right). \end{aligned} \tag{5.9}$$

Let $\mathcal{G}(m) = |h_{u,l}(m)|^2 \frac{E_S}{N_0}$ [46]. If $\mathcal{G}(1), \cdots, \mathcal{G}(m)$ are m independent exponential random variables with parameter $1/\rho$, then their sum $X = \sum_{m=1}^{g} \mathcal{G}(m)$ has an Erlang distribution. The cumulative distribution function (CDF) of X may then be expressed as [46]

$$F_X(x) = 1 - \exp\left(-\frac{x}{\rho}\right) \sum_{j=0}^{g-1} \frac{1}{j!} \left(\frac{x}{\rho}\right)^j \tag{5.10}$$

$$\begin{aligned} &= \exp\left(-\frac{x}{\rho}\right) \sum_{j=0}^{\infty} \frac{1}{j!} \left(\frac{x}{\rho}\right)^j - \exp\left(-\frac{x}{\rho}\right) \sum_{j=0}^{g-1} \frac{1}{j!} \left(\frac{x}{\rho}\right)^j \\ &= \exp\left(-\frac{x}{\rho}\right) \sum_{j=g}^{\infty} \frac{1}{j!} \left(\frac{x}{\rho}\right)^j, \end{aligned} \tag{5.11}$$

where $x = \left(2^{\Re} - 1\right)$ and g is called the shape parameter. $P_{u,l}^{\text{out,tdma}}(2)$ can then be expressed as the cdf of X in the form

$$P_{u,l}^{\text{out,tdma}}(2) = F_X\left(2^{\Re} - 1\right). \tag{5.12}$$

Note that when outage occurs for only one source node at both relays during the initial transmission, the outage probability of (5.12) can still be used to obtain the outage probability of a single source node retransmission.

Furthermore, probability that the relay R_l decodes the uth source's data packet successfully after the gth transmission is given by [41]

$$\begin{aligned} P_{u,l}^{\text{suc,tdma}}(g) &= \Pr\Big(I_{u,l}^{\text{tdma}}(g-1) < \Re, I_{u,l}^{\text{tdma}}(g) \geq \Re\Big) \\ &= \Pr\Big(I_{u,l}^{\text{tdma}}(g-1) < \Re\Big) - \Pr\Big(I_{u,l}^{\text{tdma}}(g) < \Re\Big) \\ &= P_{u,l}^{\text{out,tdma}}(g-1) - P_{u,l}^{\text{out,tdma}}(g). \end{aligned} \tag{5.13}$$

where $g = 1, 2, \ldots, G-1$. Initializing $P_{u,l}^{\text{out,tdma}}(0) = 1$, we can now summarize the probability of success for the data packet of the uth source node after the gth transmission as

$$P_{u,l}^{\text{suc,tdma}}(g) = \begin{cases} \exp\left(-\frac{\left(2^{\Re}-1\right)}{\rho}\right), & \text{for} \quad g = 1 \\ P_{u,l}^{\text{out,tdma}}(g-1) - P_{u,l}^{\text{out,tdma}}(g), & \text{for} \quad 2 \leq g \leq G-1, \end{cases} \tag{5.14}$$

assuming orthogonal retransmission.

b) *Proposed retransmission scheme*: Since the proposed retransmission scheme allows simultaneous retransmissions from the source nodes, each relay performs multi-user detection (MUD) on the received signals in order to detect the desired signal in the presence of co-channel interference (CCI).

Considering the virtual MIMO channel model given in (4.4), we introduce expressions for the signal-to-interference-plus-noise ratio (SINR) of zero forcing (ZF) and minimum mean square error (MMSE) detectors for later use. Assuming perfect knowledge of the source-relay channel at the relay R_l, the SINR for the uth source node at the ZF and MMSE detector outputs is given by [88], [89]

$$\gamma_{u,l}^{\text{zf}}(2) = \frac{\rho}{\left(\bar{\mathbf{H}}_l^{\text{H}}(2)\bar{\mathbf{H}}_l\right)_{uu}^{-1}}(2) \tag{5.15}$$

and

$$\gamma_{u,l}^{\text{mmse}}(2) = \frac{\rho}{\left(\bar{\mathbf{H}}_l^{\text{H}}(2)\bar{\mathbf{H}}_{l,2} + \rho^{-1}\mathbf{I}\right)_{uu}^{-1}} - 1, \tag{5.16}$$

respectively. The outage probability of the uth source node for the ZF detector at relay R_l is written as [89]

$$P_{u,l}^{\text{out,zf}}(2) = \Pr\Big(\log_2\left(1+\gamma_{u,l}^{\text{zf}}(2)\right) < \mathcal{R}\Big) \tag{5.17}$$

$$= F_X\left(2^{\mathcal{R}}-1\right), \tag{5.18}$$

where $F_X(x)$ is the CDF of $X \sim \chi^2_{2(N_l-N_t+1)}$ (chi-squared with $2(N_l-N_t+1)$ degrees of freedom) given in (5.10) and (5.11). The outage probability of the uth source node for the MMSE detector at relay R_l is then

$$P_{u,l}^{\text{out,mmse}}(2) = \Pr\Big(\log_2\left(1+\gamma_{u,l}^{\text{mmse}}(2)\right) < \mathcal{R}\Big). \tag{5.19}$$

In addition, the probability of success for the uth source node's data packet after two transmissions for ZF and MMSE detectors can be computed by replacing $P_{u,l}^{\text{out,tdma}}(2)$ with $P_{u,l}^{\text{out,zf}}(2)$ and $P_{u,l}^{\text{out,mmse}}(2)$, respectively, in (5.13).

5.1.2 Relay Phase Outage Probability

Now the outage probability of the two forwarding strategies during the relay phase is evaluated, under the condition that relay R_l has successfully decoded the uth source node's data packet after the gth transmission. Since, only one retransmission is allowed for each source node's data packet, the second transmission (i.e. $f=2$) during the relay phase for the uth source node's data packet is only possible if it was correctly received at R_l in its initial transmission (i.e. $g=1$).

a) *Case-I (Both relays participating in the relay phase)*: When both relays are simultaneously (re-)transmitting the data packets of their respective priority source nodes, the destination employs a MUD algorithm to recover the desired signal. The outage probability for the suboptimal linear (ZF/MMSE) detector at the destination

after the fth transmission may be expressed as

$$P_{l,D}^{\text{out,lin}}(f) = \Pr\Big(\log_2 \left(1 + \gamma_{l,D}^{\text{lin}}(f)\right) < \mathcal{R}\Big), \qquad f = 1, 2, \ldots, G-g, \tag{5.20}$$

where $\gamma_{l,D}^{\text{lin}}(f)$ denotes the SINR for relay R_l at the output of a linear detector during the fth transmission. For ZF and MMSE detectors, $\gamma_{l,D}^{\text{lin}}(f)$ is given by

$$\gamma_{l,D}^{\text{zf}}(f) = \frac{\rho}{\left(\bar{\mathbf{H}}_D^{\text{H}}(f)\bar{\mathbf{H}}_D(f)\right)_{ll}^{-1}} \tag{5.21}$$

and

$$\gamma_{l,D}^{\text{mmse}}(f) = \frac{\rho}{\left(\bar{\mathbf{H}}_D^{\text{H}}(f)\bar{\mathbf{H}}_D(f) + \rho^{-1}\mathbf{I}\right)_{ll}^{-1}} - 1, \tag{5.22}$$

respectively. Here $\bar{\mathbf{H}}_D(f) \in \mathbb{C}^{fN_d \times M}$ denotes the equivalent channel matrix during the fth transmission of the relay phase and is given by

$$\bar{\mathbf{H}}_D(f) = \begin{cases} \mathbf{H}_D(1), & \text{if } \ f = 1 \\ \begin{bmatrix} \mathbf{H}_D^{\text{T}}(1) & \mathbf{H}_D^{\text{T}}(2) \end{bmatrix}^{\text{T}}, & \text{if } \ f = 2. \end{cases} \tag{5.23}$$

When the destination employs maximum likelihood (ML) detection, the outage probability after the fth transmission can be expressed as

$$P_{l,D}^{\text{out,ml}}(f) = \Pr\left(\log_2 \det\left(\mathbf{I} + \frac{E_R}{MN_0} \bar{\mathbf{H}}_D(f)\bar{\mathbf{H}}_D^{\text{H}}(f) \right) < \mathcal{R} \right). \tag{5.24}$$

The ML detector achieves the full diversity of the channel and (5.24) can be taken as a lower bound on outage probability performance.

Note that if the mutual information for only one relay-destination link falls below the target rate during the initial transmission (i.e. $f = 1$) of the relay phase, then only the relay that is in outage will participate in the relay phase by retransmitting its information in the next time slot. The outage probability after the second transmission (i.e. $f = 2$) in such a case can be obtained by using the results of (5.28) given below.

b) *Case-II (Single relay participating in the relay phase)*: If only one active relay is available to forward the data packet of source node(s) over orthogonal channels, then the mutual information for the link between the lth relay and the destination in the fth transmission may be written as

$$I_{l,D}^{\text{tdma}}(f) = \log_2\left(1 + \sum_{n=1}^{f} \|\mathbf{h}_l(n)\|_F^2 \frac{E_R}{N_0}\right), \qquad f = 1, 2, \ldots, G-g. \tag{5.25}$$

The outage probability after the fth transmission is then

$$\begin{aligned} P_{l,D}^{\text{out,tdma}}(f) &= \Pr\left(I_{l,D}^{\text{tdma}}(f) < \mathcal{R}\right) \\ &= \Pr\left(\sum_{n=1}^{f} \|\mathbf{h}_l(n)\|_F^2 \frac{E_R}{N_0} < \left(2^{\mathcal{R}} - 1\right)\right) \qquad (5.26) \\ &= F_X(x), \qquad (5.27) \end{aligned}$$

where $x = \left(2^{\mathcal{R}} - 1\right)$. We let $Z = fN_d$. Since $\|\mathbf{h}_l(n)\|_F^2$ has a Chi-square distribution with $2N_d$ degrees of freedom, the outage probability can be expressed as the CDF of X given in (5.10), as

$$P_{l,D}^{\text{out,tdma}}(f) = 1 - \exp\left(-\frac{x}{\rho}\right) \sum_{j=0}^{Z} \frac{1}{j!} \left(\frac{x}{\rho}\right)^j. \tag{5.28}$$

5.1.3 End-to-End (e2e) Outage Probability

Finally, the end-to-end outage probability of the two-hop HARQ DF relaying system is analyzed. The e2e outage probability of the uth source node's data packet at the destination after G transmissions can be expressed as [46]

$$P_{u,G}^{\text{out,e2e}} = \sum_{g=1}^{G-1} P_{u,l}^{\text{suc,mode}}(g) P_{l,D}^{\text{out,mode}}(G-g) + P_{u,l}^{\text{out,mode}}(G-1). \tag{5.29}$$

Depending upon the (re-)transmission scheme and the MUD algorithm employed at the receiver, the "mode"can be chosen accordingly in order to evaluate the e2e outage

probability.

5.2 Average Throughput Analysis

In this section, the throughput performance of the proposed two-hop HARQ DF relaying system is evaluated. First, the average throughputs of orthogonal (time division) retransmission and proposed retransmission schemes during the broadcast phase are evaluated. Then an expression for the average end-to-end throughput performance for the two schemes is developed.

The system throughput efficiency can be obtained using the renewal reward theorem [48, 90], where a system renews itself after every packet transmission [91]. The average duration of a single renewal period is equal to the length of the packet. The idea behind the renewal reward process is to associate a reward with every packet transmission [91]. If a packet contains p bits, then the reward is p if all the bits are successfully decoded at the receiver and 0 if at least one bit contains error [91]. Using the renewal-reward theorem [48], the average throughput for node a at node b with the maximum number of transmissions (e.g. J) can be written as

$$\eta_{a,b,J} = \frac{E\{\Psi_{a,b,J}\}}{E\{\mathcal{T}_{a,b,J}\}}, \tag{5.30}$$

where $E\{\Psi_{a,b,J}\}$ is the average reward (the expected number of successfully decoded bits of node a at node b) and $E\{\mathcal{T}_{a,b,J}\}$ is the average transmission time (the expected number of time slots used by node a for the initial transmission and retransmissions). Initializing $P_{a,b}^{\text{out}}(0) = 1$, the average reward $\Psi_{a,b,J}$ and average time slots $\mathcal{T}_{a,b,J}$ are

given by

$$\begin{aligned} E\{\Psi_{a,b,J}\} &= \Re\sum_{j=1}^{J}\Big(P_{a,b}^{\text{out}}(j-1)-P_{a,b}^{\text{out}}(j)\Big) \\ &= \Re\Big(1-P_{a,b,J}^{\text{out}}\Big), \end{aligned} \tag{5.31}$$

$$\begin{aligned} E\{\mathcal{T}_{a,b,J}\} &= \sum_{j=1}^{J} j\Big(P_{a,b}^{\text{out}}(j-1)-P_{a,b}^{\text{out}}(j)\Big)+JP_{a,b,J}^{\text{out}} \\ &= \sum_{j=0}^{J-1} P_{a,b}^{\text{out}}(j) \\ &= 1+\sum_{j=1}^{J-1} P_{a,b}^{\text{out}}(j), \end{aligned} \tag{5.32}$$

respectively. Substituting (5.31) and (5.32) into (5.30) yields the throughput as

$$\eta_{a,b,J}=\frac{\Re\left(1-P_{a,b,J}^{\text{out}}\right)}{1+\sum_{j=1}^{J-1}P_{a,b}^{\text{out}}(j)}. \tag{5.33}$$

5.2.1 Broadcast Phase Throughput

Here, we derive average throughput expressions for the orthogonal (time division) and proposed retransmission scheme during the broadcast phase.

a) *Orthogonal (time division) retransmission scheme*: Using the results of (5.33), the average total reward (the sum of each individual source node's average reward) for the orthogonal (time division) retransmission scheme at relay R_l after the second transmission ($g=2$) can be expressed as

$$\begin{aligned} E\{\Psi^{\text{orth}}(2)\} &= \sum_{u=1}^{2} E\{\Psi_{u,l}^{\text{orth}}(2)\} \\ &= \sum_{u=1}^{2}\Re\left(1-P_{u,l}^{\text{out,tdma}}(2)\right), \end{aligned} \tag{5.34}$$

where $P_{u,l}^{\text{out,tdma}}(2)$ is the uth source node outage probability at relay R_l after the second transmission as given in (5.12). The average total number of time slots used by both

source nodes for the orthogonal (time division) retransmission scheme is given by

$$\begin{aligned} E\{\widehat{\mathcal{T}}^{\text{orth}}(2)\} &= \sum_{u=1}^{2}\left(1+P_{u,l}^{\text{out,tdma}}(1)\right) \\ &= \hat{\tau}_{\text{int}}^{\text{tdma}}+\sum_{u=1}^{2}P_{u,l}^{\text{out,tdma}}(1), \end{aligned} \tag{5.35}$$

where $\hat{\tau}_{\text{int}}^{\text{tdma}}$ is the transmission time required for both source nodes during the initial transmission of the broadcast phase. As discussed in Section 4.3.1, both source nodes transmit their data packets over two successive time slots during the initial transmission, therefore $\hat{\tau}_{\text{int}}^{\text{tdma}}=2$. The total throughput for the orthogonal (time division) retransmission scheme can then be expressed as

$$\eta^{\text{orth}}(2)=\frac{\sum_{u=1}^{2}\Re\left(1-P_{u,l}^{\text{out,tdma}}(2)\right)}{2+\sum_{u=1}^{2}P_{u,l}^{\text{out,tdma}}(1)}. \tag{5.36}$$

b) *Proposed retransmission scheme*: The average total reward for the proposed retransmission scheme after the second transmission is given by

$$E\{\Psi^{\text{prop}}(2)\}=\sum_{u=1}^{2}\Re\left(1-P_{u,l}^{\text{out,lin}}(2)\right), \tag{5.37}$$

where $P_{u,l}^{\text{out,lin}}(2)$ is the outage probability of the uth source node for a suboptimal linear (ZF/MMSE) detector at relay R_l given in (5.17) or (5.19). In the proposed scheme, both source nodes simultaneously retransmit their data packets within a single time slot. Since both source nodes transmit their data packets in time division fashion, the number of time slots required during the second transmission of the proposed retransmission scheme can be found as

$$\hat{\tau}^{\text{prop}}(2)=\max\left(P_{1,l}^{\text{out,tdma}}(1),P_{2,l}^{\text{out,tdma}}(1)\right). \tag{5.38}$$

Then, the average total number of time slots used by both source nodes during the second transmission is the sum of the number of time slots used for the time division

initial transmissions and simultaneous retransmission and can be written as

$$E\{\widehat{\mathcal{T}}^{\text{prop}}(2)\} = 2 + \max\left(P_{1,l}^{\text{out,tdma}}(1), P_{2,l}^{\text{out,tdma}}(1)\right). \tag{5.39}$$

Comparing (5.35) and (5.39), we can see that when outage occurs for both source nodes during the initial transmission, orthogonal (time division) retransmission requires four time slots to complete the second transmission. On the other hand, the proposed scheme uses only three time slots to complete the second transmission. The overall throughput for the proposed retransmission scheme is then

$$\eta^{\text{prop}}(2) = \frac{\sum_{u=1}^{2} \mathcal{R}\left(1 - P_{u,l}^{\text{out,lin}}(2)\right)}{2 + \max\left(P_{1,l}^{\text{out,tdma}}(1), P_{2,l}^{\text{out,tdma}}(1)\right)}. \tag{5.40}$$

5.2.2 End-to-End (e2e) Throughput

Finally, we derive the e2e throughput expressions for the two-hop HARQ DF relaying system. The average total e2e reward at the destination after G transmissions is written as

$$E\{\Psi_G^{\text{out,e2e}}\} = \sum_{u=1}^{2} \mathcal{R}\left(1 - P_{u,G}^{\text{out,e2e}}\right), \tag{5.41}$$

where $P_{u,G}^{\text{out,e2e}}$ is the e2e outage probability of the uth source node at the destination given in (5.29). The average total e2e transmission time required for both source nodes' data packets for the proposed two-hop HARQ DF relaying system can be expressed as [46]

$$E\{\mathcal{T}_G^{\text{e2e}}\} = E\{\widehat{\mathcal{T}}^{\text{mode}}(G-1)\} + E\{\widetilde{\mathcal{T}}^{\text{mode}}(G-g)\}. \tag{5.42}$$

Then the total e2e throughput after G transmissions is given by

$$\eta_G^{\text{e2e}} = \frac{E\{\Psi_G^{\text{out,e2e}}\}}{E\{\mathcal{T}_G^{\text{e2e}}\}}. \tag{5.43}$$

In (5.42), the first term on the right-hand side represents the average number of time slots used for both source nodes' data packets during $(G-1)$ transmissions of the

broadcast phase. Letting $G = 3$, the average number of time slots required for the orthogonal (time division) and proposed retransmission schemes are given in (5.35) and (5.39), respectively. The second term on the right-hand side represents the average number of time slots used for forwarding both source nodes' data packets during $(G-g)$ transmissions of the relay phase. Under the condition that relay R_l has successfully decoded the packet of the uth source node after the gth transmission, the average transmission time required for the two forwarding schemes during $(G-g)$ transmissions can be evaluated.

a) *Case-I (Both relays participating in the relay phase)*: Under this forwarding strategy, both relays will simultaneously (re-)transmit the data packets of their respective priority source nodes. For simplicity and better understanding of the throughput analysis, we assume that the relays R_1 and R_2 always forward the data packet's of S_1 and S_2, respectively. The average number of time slots required for case-I during $(G-g)$ transmissions is expressed as

$$E\{\widetilde{\mathcal{T}}^{\text{case-I}}(G-g)\} = \sum_{g=1}^{G-1} \max\left(P_{1,l}^{\text{succ,mode}}(g) \sum_{f=0}^{G-g-1} P_{1,D}^{\text{out,mud}}(f), P_{2,l}^{\text{succ,mode}}(g) \sum_{f=0}^{G-g-1} P_{2,D}^{\text{out,mud}}(f) \right), \tag{5.44}$$

where $P_{l,D}^{\text{out,mud}}(f)$ is the outage probability of the uth source node (forwarded by the relay R_l) for the MUD algorithm at the destination as specified in (5.20) and (5.24). We note that the average number of time slots used for case-I during the relay phase is determined by the maximum of $\sum_{f=0}^{G-g-1} P_{1,D}^{\text{out,mode}}(f)$ and $\sum_{f=0}^{G-g-1} P_{2,D}^{\text{out,mode}}(f)$. For $l = 1, 2$, the term, $\sum_{f=0}^{G-g-1} P_{l,D}^{\text{out,mode}}(f)$ represents the average number of time slots used during $(G-g)$ transmissions of the relay phase, when the relay R_l successfully decodes the data packet of the uth source node after g transmissions, where $g = 1, 2, \ldots, G-1$.

b) *Case-II (Single relay participating in the relay phase)*: If relay R_l successfully decodes the data packets of both source nodes during the broadcast phase, then the average total number of time slots used for case-II during $(G-g)$ transmissions is the

sum of the number of time slots used by relay R_l to forward the data packet of each source node. It is given by

$$E\{\widetilde{\mathcal{T}}^{\text{case-II}}(G-g)\} = \sum_{g=1}^{G-1}\sum_{u=1}^{2}\left(P_{u,l}^{\text{succ,mode}}(g)\sum_{f=0}^{G-g-1}P_{l,D}^{\text{out,tdma}}(f)\right), \tag{5.45}$$

where $P_{l,D}^{\text{out,tdma}}(f)$ is the outage probability of the uth source node (forwarded by the relay R_l) for case-II after the fth transmission as given in (5.28).

5.3 Numerical and Simulation Results

Here, the outage and throughput performance for the proposed two-hop HARQ DF relaying system is evaluated and compared with simulation results, obtained by averaging over 10^6 channel realizations. Each source and relay node is equipped with one antenna. The destination is equipped with $N_d = 2$ receive antennas. A quasi-static Rayleigh frequency-flat fading channel is assumed for all links. The maximum number of retransmissions allowed for each source node is one. For all schemes, we consider a target transmission rate of $\mathcal{R} = 2$ (bits/s/Hz).

5.3.1 Outage Probability Results

Figure 5.1 shows outage probability performance for the broadcast phase. As is evident in this figure, the outage probability using orthogonal (time division) transmissions, given in (5.7) and (5.12), are almost identical to the Monte-Carlo simulation results for the same number of transmissions. Moreover, it is observed that the diversity order of an orthogonal (time division) transmission scheme increases as the number of HARQ rounds increases. However, increasing the number of HARQ rounds comes with a loss in throughput. Furthermore, we see that the orthogonal (time division) retransmission scheme performs better than the proposed retransmission scheme. This is due to the co-channel interference during simultaneous retransmission and also due to the fact that the relay is employing a suboptimal linear MUD technique.

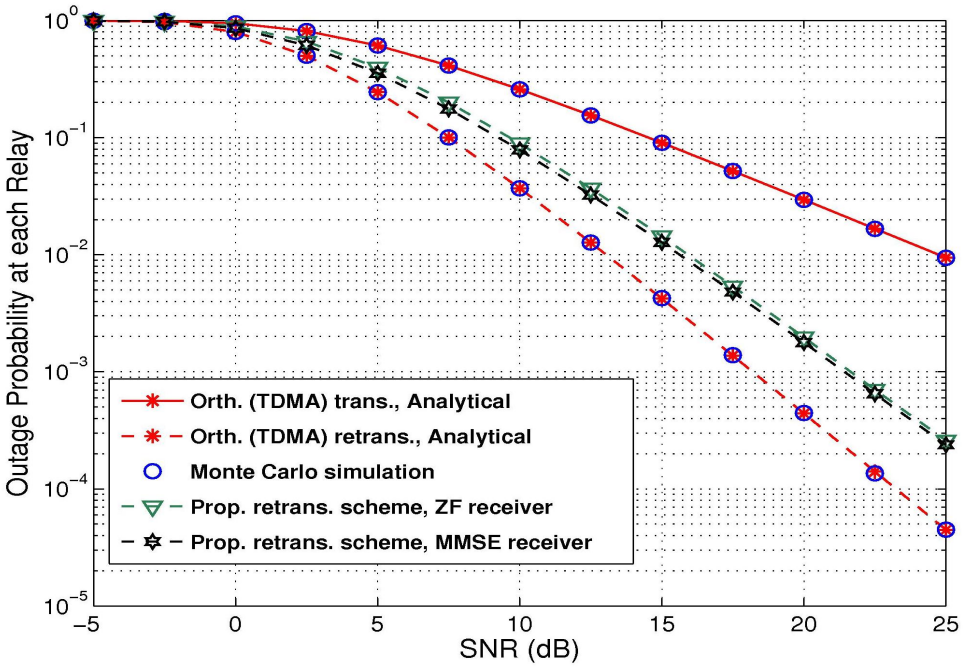


Figure 5.1: Outage performance of sources-relays (broadcast) link. First transmission (solid lines), second transmission (dashed lines) and $\mathcal{R} = 2$ (bits/s/Hz).

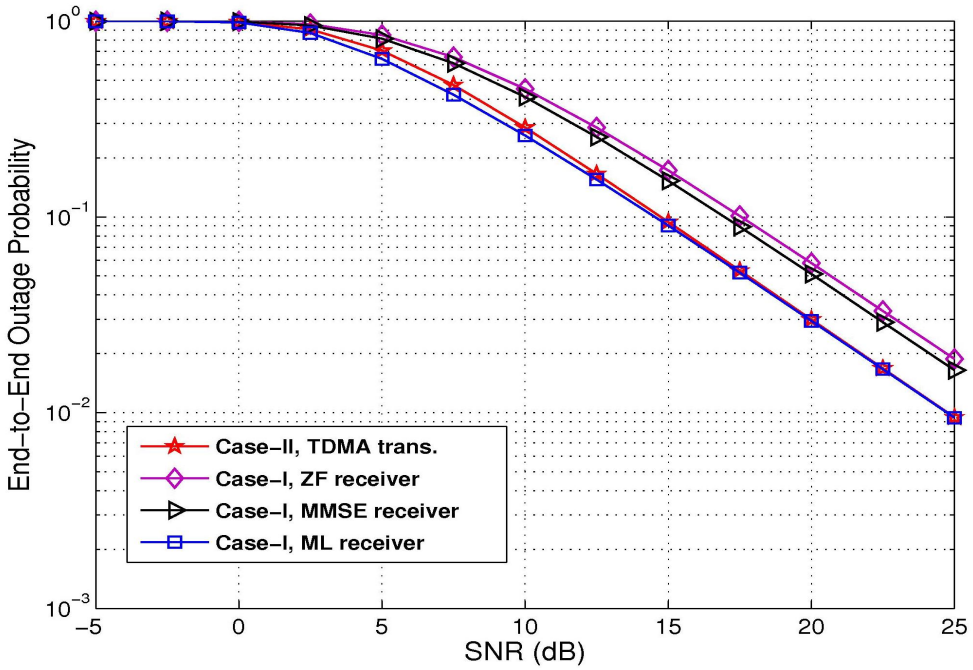


Figure 5.2: End-to-end outage performance for the two forwarding strategies of a two-hop DF relaying system with $G = 2$ (no HARQ round) and $\mathcal{R} = 2$ (bits/s/Hz).

With a maximum, $G = 2$, transmissions (no HARQ round for each source node's data packet), the end-to-end outage probability of a two-hop DF relaying system for the two forwarding cases is shown in Figure 5.2. As we can see, in a low SNR region Case-I (both relays simultaneously forwarding a different source node's data packets) with a ML receiver achieves marginal outage performance gain over Case-II (single relay forwarding in different time slots). For high SNR both cases yield similar outage performance. This is because the initial transmission during the broadcast phase is carried out in time division fashion and in general, the e2e outage probability performance of a two-hop DF relaying system with no direct path is limited by the outage performance of any link, i.e. broadcast or relay link. The e2e outage probability of Case-I with suboptimal linear receivers is also shown in Figure 5.2.

Figures 5.3 and 5.4 illustrate the e2e outage performance for the two forwarding strategies with a maximum number of transmissions, $G = 3$ (single retransmission for each source node's data packet), while considering both orthogonal (time division) retransmission and the simultaneous (proposed) retransmission schemes, respectively, during the broadcast phase. Comparing the orthogonal retransmissions (Figure 5.3) and the proposed retransmissions (Figure 5.4) with MMSE MUD employed at each relay, it is observed that at high SNR the e2e outage performance slightly improves for both cases, i.e. Case-I and Case-II, when orthogonal (time division) retransmissions are considered instead of simultaneous retransmissions during the broadcast phase. Moreover, from Figure 5.4 it is noted that the e2e outage performance gap between Case-I with ML receiver and Case-II decreases, showing similar performance at high SNR. On the other hand Case-I with MMSE receiver has performance slightly worse than Case-II. For example, at outage probability of 10^{-3}, there is a 0.6 dB performance gap between Case-II and Case-I for a MMSE receiver. This is for the same reason as in Figure 5.2.

The outage probability results for the proposed two-hop HARQ DF relaying system suggest that the outage performance for each link and the e2e outage performance sig-

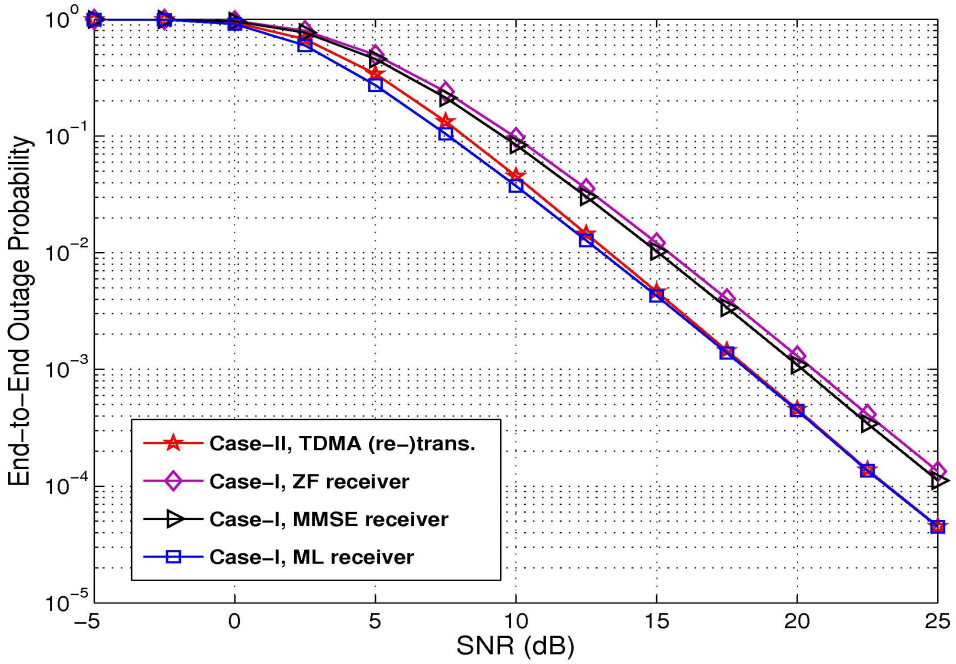


Figure 5.3: End-to-end outage performance for the two forwarding strategies of a two-hop DF relaying system with orthogonal (time division) retransmissions during the broadcast link. $G = 3$ (one HARQ round) and $\mathcal{R} = 2$ (bits/s/Hz).

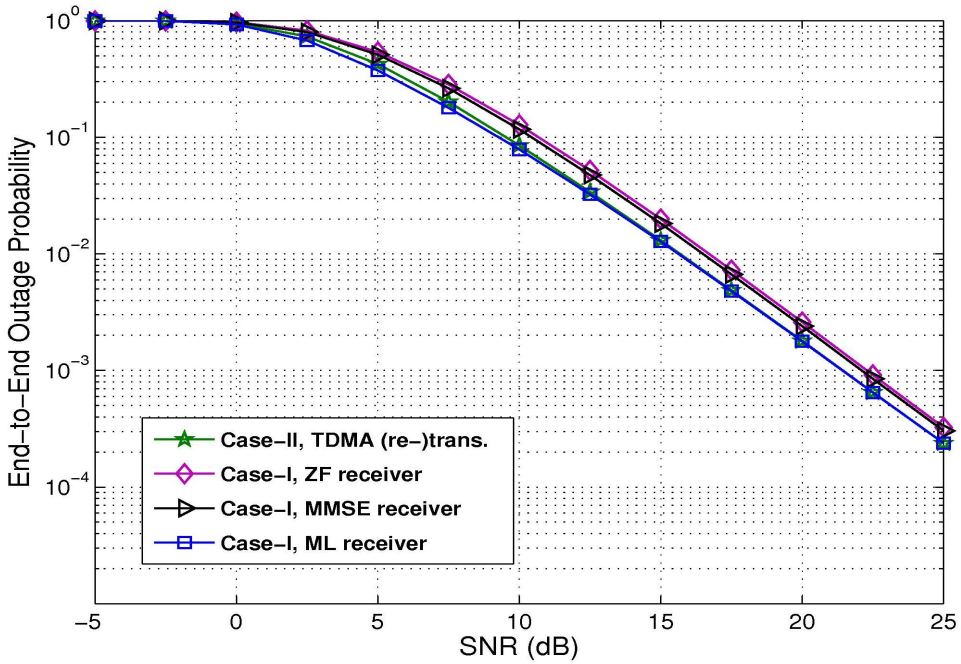


Figure 5.4: End-to-end outage performance for the two forwarding strategies of a two-hop DF relaying system with simultaneous (proposed) retransmissions during the broadcast link and MMSE MUD employed at each relay. $G = 3$ (one HARQ round) and $\mathcal{R} = 2$ (bits/s/Hz).

nificantly improve if both the relays and the destination employ ML MUD during the (re-)transmission round. Unfortunately, the complexity of ML receivers grows exponentially with the number of transmitted signals, making it infeasible for many practical applications.

5.3.2 Throughput Results

Figure 5.5 shows the average throughput performance of the orthogonal (time division) retransmission and simultaneous (proposed) retransmission scheme during the broadcast phase. From the figure, it is observed that at low SNR both retransmission schemes yield similar throughput performance. For SNR greater than 5dB, we note that the proposed retransmission scheme achieves higher throughput than the orthogonal (time division) retransmission scheme. As the SNR increases, both retransmission schemes have a maximum achievable throughput for a given rate, $\mathcal{R}$. This is due to the fact that when the SNR is high enough, retransmissions from both source nodes will rarely happen.

Figure 5.6 compares the average e2e throughput performance for the two forwarding strategies with a maximum number of transmissions, $G = 2$ (no HARQ round for each source node's data packet). As we can see, in the medium and high SNR regions, Case-I achieves higher throughput than Case-II. However, when the destination employs suboptimum linear MUD algorithms (ZF and MMSE), Case-II yields better throughput performance than Case-I (ZF and MMSE) at medium SNR. In addition, it is observed that at high SNR, the average e2e throughput performance of Case-I (for all MUD algorithms) and case-II will floor at 1.32 and 1 (bits/s/Hz), respectively. This is due to the fact that in the high SNR region, at least three and four time slots are required to transmit both source node's data packets to the destination for Case-I and Case-II, respectively.

Figures 5.7 and 5.8 show the average e2e throughput performance for the two forwarding strategies with a maximum number of transmissions, $G = 3$ (single retrans-

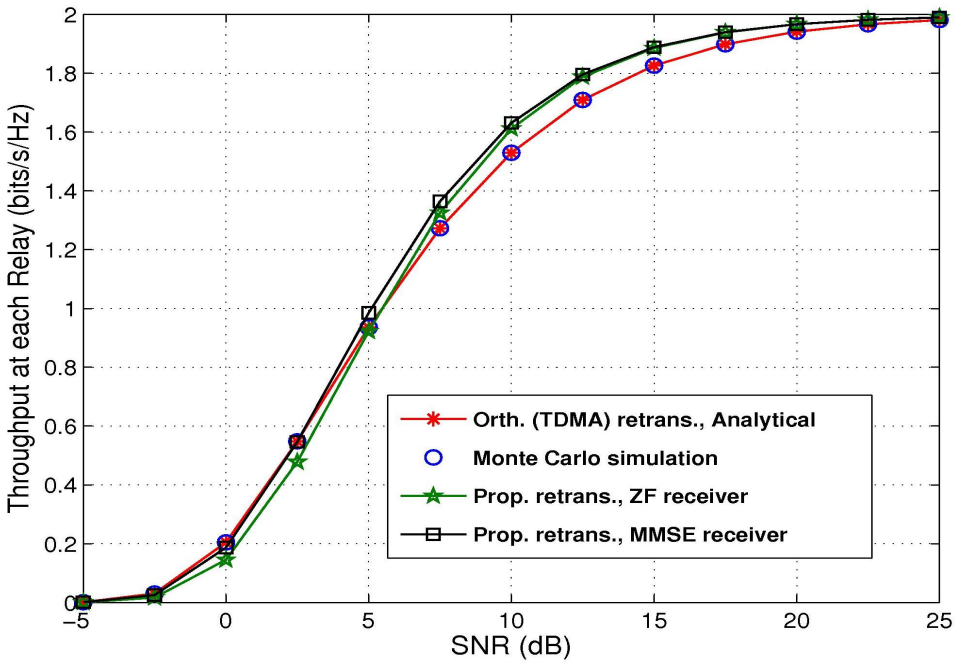


Figure 5.5: Throughput performance for orthogonal (time division) and simultaneous (proposed) retransmission schemes during the broadcast phase. $\mathcal{R} = 2$ (bits/s/Hz).

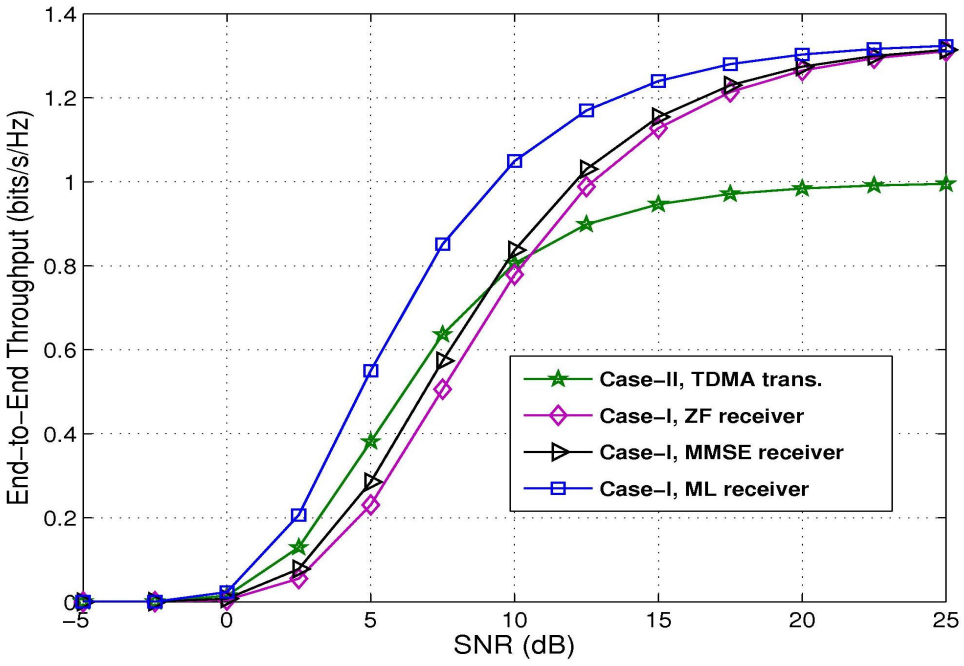


Figure 5.6: End-to-end throughput performance for the two forwarding strategies of a two-hop DF relaying system with $G = 2$ (no HARQ round) and $\mathcal{R} = 2$ (bits/s/Hz).

mission for each source node's data packet), while considering both orthogonal (time division) retransmission and the simultaneous (proposed) retransmission schemes, respectively, during the broadcast phase. Comparing the two figures, it is noted that the average e2e throughput performance for both cases is improved when the simultaneous (proposed) retransmission scheme is considered instead of the orthogonal (time division) retransmission scheme during the broadcast phase. Furthermore, we can see that the average e2e throughput performance of Case-I (for all MUD algorithms) and Case-II floors at 1.32 and 1 (bits/s/Hz), respectively in the high SNR region. This is for the same reason as in Figure 5.6.

Comparing Figures 5.6-5.8, it is observed that in the low and medium SNR region, the average e2e throughput performance increases when we increase the total number of transmissions, G, for each source node's data packet.

5.4 Chapter Conclusions

In this chapter, the system outage probability was analyzed for various transmission modes and it was observed that orthogonal (time division) (re-)transmissions yields better outage performance than the proposed (simultaneous) (re-)transmission schemes with suboptimal linear MUD algorithms. However, time division (re-)transmission scheme requires more time slots to complete the HARQ round than the proposed (re-)transmission scheme. As a result a time division (re-)transmission scheme can degrade the system throughput, especially the end-to-end throughput of a DF relaying system. It has also been shown that the outage performance for the proposed schemes can further be improved if the receiver employs a ML MUD algorithm. The renewal-reward theorem was used to derive throughput expressions for the proposed DF relaying system. The throughput performance of time division (re-)transmissions was compared with the proposed schemes and it was found that the proposed (re-)transmission schemes achieve higher throughput than the traditional time division (re-)transmissions.

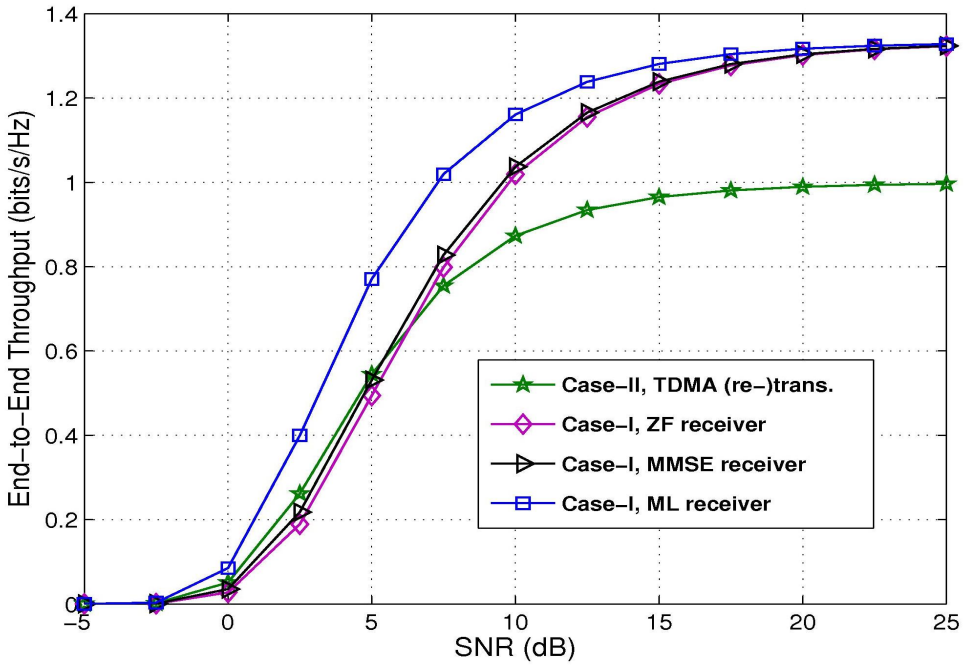


Figure 5.7: End-to-end throughput performance for the two forwarding strategies of a two-hop DF relaying system with orthogonal (time division) retransmissions during the broadcast link. $G = 3$ (one HARQ round) and $\mathcal{R} = 2$ (bits/s/Hz).

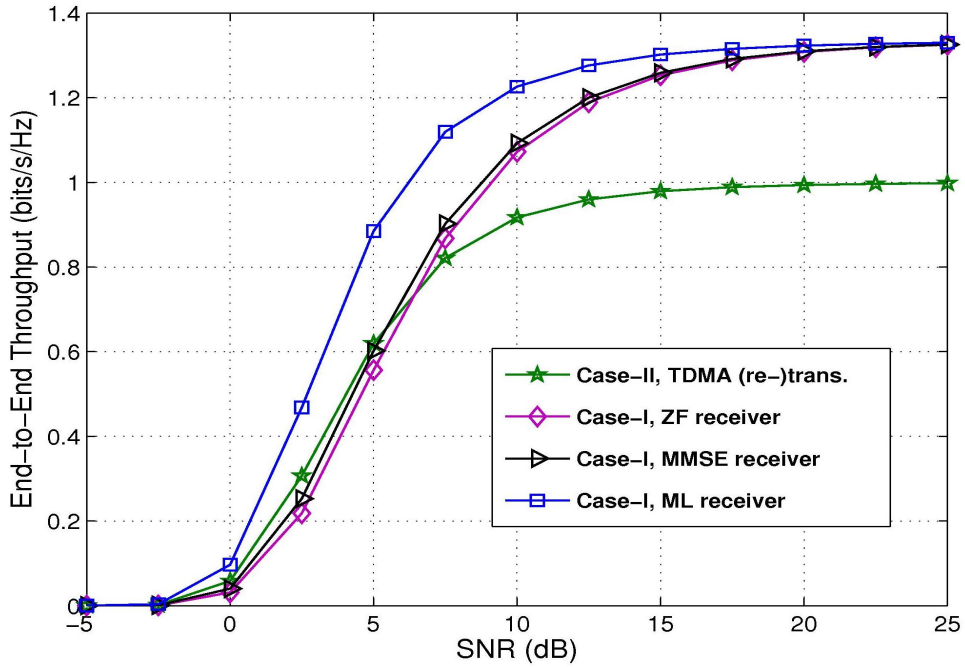


Figure 5.8: End-to-end throughput performance for the two forwarding strategies of a two-hop DF relaying system with simultaneous (proposed) retransmissions during the broadcast link and MMSE MUD employed at each relay. $G = 3$ (one HARQ round) and $\mathcal{R} = 2$ (bits/s/Hz).

Chapter 6

Conclusions and Future Work

This chapter summarizes the contributions of the thesis and highlights several potential extensions and open problems for future research.

6.1 Summary and Conclusions

This thesis focused on several design aspects of multiple HARQ process over two different wireless uplink systems: the non-cooperative system (Chapter 3) and the cooperative relaying system (Chapter 4 & 5). Throughout the thesis, it is assumed that the wireless channels exhibit quasi-static Rayleigh frequency-flat fading and that perfect knowledge of CSI is available at the receiver. The research work presented in this thesis is summarized as follows:

Typically the suboptimal linear MUD algorithms, such as ZF and MMSE fail to perform under overload and ML detection is often prohibitively complex. Motivated by this drawback, a simple and novel multi-user transmission approach that combines the HARQ retransmissions with the use of virtual receive antennas [23] was presented in Chapter 3. The main idea was to benefit not only from the HARQ retransmissions, but also to treat the HARQ retransmissions as if they were arriving to an additional virtual receive antennas. The proposed approach allowed us to apply linear MUD algorithms to perform joint HARQ detection on the stacked retransmissions without requiring additional antennas or hardware chains.

Based on the proposed idea, two detection schemes were described in Chapter 3. In scheme-I, the user with a successfully decoded packet was kept silent and retransmission was performed by the user with an erroneous packet. In scheme-II, the user with a successfully decoded packet was allowed to transmit new information instead of remaining idle. The performance of the proposed schemes was evaluated in terms of bit-error-rate (BER), dropped packet-rate ($\mathcal{P}_{rate}$) and throughput efficiency (η). The simulation results demonstrated that:

i. The proposed schemes allow us to transform overloaded conditions to critically loaded conditions by allowing sufficient transmissions.

ii. For the same number of transmissions, scheme-I achieves better BER and $\mathcal{P}_{rate}$ performance than scheme-II for both ML and MMSE MUD techniques due to the decrease in CCI following each (re-)transmission.

iii. For the same number of transmissions, scheme-II achieves marginal throughput performance gain over scheme-I.

In Chapter 4, a two-hop DF relaying system with HARQ protocols was studied, where two source nodes communicate with a single destination with the assistance of two relay nodes. We focused on the scenario in which the direct link between the source nodes and the destination does not exist due to large physical separation and path loss. A multi process HARQ retransmission scheme was proposed to increase the throughput efficiency by allowing simultaneous retransmissions from both source nodes only when necessary. After completion of each (re-)transmission round during the broadcast phase, relays were allowed to exchange their decoding outcomes among themselves in order to avoid unnecessary retransmissions. Simulation results showed that exchanging decoding outcomes among relay nodes resulted in improved average retransmission rate. Finally, a novel forwarding strategy was proposed to maximize the number of simultaneous (re-)transmissions during the relay phase.

In Chapter 5, the performance of the two-hop DF relaying system introduced in

Chapter 4 was analyzed. First the outage probabilities of each transmission phase (i.e. the broadcast phase and the relay phase) and then the end-to-end outage probabilities were evaluated for both the simultaneous and orthogonal time division (re-)transmissions. Finally, the throughput efficiency of the proposed retransmission and forwarding schemes was analyzed using the renewal-reward theorem. For comparison purposes, the throughput efficiency of orthogonal time division (re-)transmissions was also evaluated. The analytical results were found to be almost identical to the simulation results. The analytical results demonstrated that:

i. Orthogonal (time division) retransmission scheme yields better performance in terms of outage probability than the proposed retransmission scheme.

ii. The e2e outage probability performance of a two-hop DF relaying system in the absence of a direct path is limited by the outage performance of either the broadcast or relay link.

iii. The proposed (re-)transmission schemes achieve higher throughput than the traditional orthogonal (time division) (re-)transmissions.

6.2 Future Work

Several possible extensions of the research work presented in this thesis are outlined as follow:

- Throughout the thesis, it is assumed that perfect channel state information (CSI) is available at the receiver and that the feedback channel is error free with negligible delay. None of these assumptions are realistic, although they are commonly employed in the literature. Therefore, it would be interested to carry out an in-depth study of multi process HARQ transmissions with imperfect CSI and feedback channel. The derivations in Chapter 5 can be further extended to include the effects of corrupted and delayed feedback channel for HARQ retransmissions [92].

- Multiple HARQ transmission schemes were studied over frequency-flat fading channels. The extension to frequency-selective fading channels is another potential prospect of this work. A frequency-selective fading channel may be represented as a combination of several narrowband channels shifted in the delay domain. It is more complicated than the frequency-flat fading channel. Therefore, designing low-complexity detection approaches for the proposed HARQ transmissions over frequency-selective fading channels would be an interesting topic to address in future.

- For the two-hop DF relaying system (Chapter 4 & 5), we have not investigated the power allocation between the source nodes and the relay nodes to evaluate the outage and throughput performance. Recently, adaptive and optimal power allocation among the nodes is being studied extensively to minimize the outage probability under a total average transmit power constraint [93, 94, 95]. Since the overall performance of a cooperative relaying system largely depends on power allocation schemes [96], therefore designing efficient power allocation scheme can be of great interest in order to improve the system performance. An interesting possibility is to use the HARQ feedback signal to do some form of power allocation.

- The throughput efficiency of multi-source and multi-relay cooperative network can be degraded, if the intermediate nodes operates on orthogonal (time division) channels. In order to improve the transmission efficiency for such networks, the relay node can combine the signals received from multiple source nodes through network coding (NC) and forward a single mixed (i.e. network coded) signal to the destination(s) rather than forwarding the individual signals separately [97]. Recently, it has been shown that the combination of network coding and HARQ protocols can result in improved throughput efficiency over traditional HARQ schemes [94, 98]. An interesting question that could be investigated in future is: how to use network coding to further improve the performance of our proposed

multi process HARQ schemes?

- The HARQ retransmission and relay forwarding schemes proposed and analyzed in Chapter 4 and Chapter 5 considered a two-source and two-relay DF relaying network. Based on the average throughput numerical results, it can be concluded that extending the proposed retransmission and forwarding schemes for a system with a large number of source and relay nodes can further improve the throughput efficiency of a two-hop DF relaying system compared to orthogonal (re-)transmissions. However, this is not a straightforward extension. Applying the proposed forwarding scheme in a DF relaying network with more than two sources and relays becomes a non-trivial task. Therefore, it is important to develop an efficient and simple forwarding scheme for large relay networks that treats all source nodes fairly and does not require large amounts of extra resources.

References

[1] M. Uthansakul and P. Uthansakul, "Experiments with a low-profile beamforming MIMO system for WLAN applications," *IEEE Antennas and Propagation Mag.*, vol. 53, no. 6, pp. 56 –69, Dec. 2011. 2

[2] T. Siadari, F. Rezha, and S.-Y. Shin, "Applied MIMO based on Hadamard transform in IEEE 802.16 WiMAX system," in *Proc. IET Int'l. Conf. on Inform. and Commun. Tech. (IETICT)*, 27-29 Apr. 2013, pp. 582 – 586. 2

[3] B. Badic, R. Balraj, T. Scholand, Z. Bai, and S. Iwelski, "Impact of feedback and user pairing schemes on receiver performance in MU-MIMO LTE systems," in *Proc. IEEE Wireless Commun. and Networking Conf. (WCNC)*, 1-4 Apr. 2012, pp. 399–403. 2

[4] S. Yang and J.-C. Belfiore, "Diversity-multiplexing tradeoff of double scattering MIMO channels," *IEEE Trans. Inform. Theory*, vol. 57, no. 4, pp. 2027 –2034, Apr. 2011. 2

[5] T.-W. Tang, M.-K. Chen, and H.-F. Lu, "Improving the DMT performance for MIMO communication with linear receivers," *IEEE Trans. Veh. Tech.*, vol. 62, no. 3, pp. 1189 –1200, Mar. 2013. 2

[6] L. Badia, P. Casari, M. Levorato, and M. Zorzi, "Analysis of an automatic repeat request scheme addressing long delay channels," in *Proc. IEEE Int'l. Conf. on Advanced Info. Networking and Applications Workshops (WAINA)*, 26-29 May 2009, pp. 1142 –1147. 3, 31

[7] S. Lin and D. J. Costello, Jr., *Error Control Coding*, 2nd ed. Prentice Hall, 2004. 3, 31, 32, 33

[8] D. Chase, "Code combining- a maximum likelihood decoding approach for combining an

arbitrary number of noisy packets," *IEEE Trans. Commun.*, vol. 33, no. 5, pp. 385 –393, May 1985. 3, 34

[9] Y. Wu and S. Xu, "Energy-efficient multi-user resource management with IR-HARQ," in *Proc. IEEE Veh. Tech. Conf. (VTC)*, 6 - 9 May 2012, pp. 1 –5. 3, 34

[10] T. Lestable and M. Ran, *Error Control Codining for B3G/4G Wireless Systems: Paving the Way to IMT-Advanced Standards*, 1st ed. John Wiley & Sons, 2011. 4, 32, 33, 36, 37

[11] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs. Tech. J.*, vol. 1, no. 2, pp. 41 –59, 1996. 4

[12] H.-Y. Chen, C.-H. Chuang, P.-C. Yeh, and S. Zummo, "Optimal power allocation and power control for VBLAST systems with M-ary modulations," *IET Commun.*, vol. 4, no. 8, pp. 956 –966, May 2010. 4

[13] L. Bai and J. Choi, *Low Complexity MIMO Detection*. Springer, 2012. 5, 27, 28

[14] S. Zhang, C. Nie, L. Lu, S. Zhang, and G. Qian, "MIMO physical layer network coding based on VBLAST detection," in *Proc. IEEE Int'l. Conf. on Wireless Commun. & Sig. Process. (WCSP)*, 25 -27 Oct. 2012, pp. 1–5. 5, 46

[15] B. W. Zarikoff, J. K. Cavers, and S. Bavarian, "An iterative groupwise multiuser detector for overloaded MIMO applications," *IEEE Trans. Wireless Commun.*, vol. 6, no. 2, pp. 443 –447, Feb. 2007. 5, 46, 47

[16] J. Lee, D. Toumpakaris, and W. Yu, "Interference mitigation via joint detection," *IEEE J. Select. Areas Commun.*, vol. 29, no. 6, pp. 1172 –1184, Jun. 2011. 5

[17] M. Krause, D. P. Taylor, and P. A. Martin, "List-based group-wise symbol detection for multiple signal communications," *IEEE Trans. Wireless Commun.*, vol. 10, no. 5, pp. 1636 –1644, May 2011. 5, 47

[18] M. O. Damen, K. Abed-Meraim, and J.-C. Belfiore, "A generalised sphere decoder for asymmetrical spacetime communication architecture," *IEEE Electron. Lett.*, vol. 36, no. 2, pp. 166 –167, Jan. 2000. 5

[19] K. T. Tasneem, P. A. Martin, and D. P. Taylor, "Iterative soft detection of cochannel signals using ant colony optimization," in *Proc. IEEE Int'l. Symp. on Personal, Indoor and Mobile Radio Commun. (PIMRC)*, 9-12 Sep. 2012, pp. 1617 –1621. 5, 47

[20] K. Obaidullah and Y. Miyanaga, "Efficient algorithm with lognormal distributions for overloaded MIMO wireless system," in *Proc. IEEE Sig. & Info. Process. Assoc. Annual Summit and Conf. (APSIPA ASC), Asia-Pacific*, 3 -6 Dec. 2012, pp. 1–4. 5, 47

[21] Z. Rauf, P. A. Martin, and D. P. Taylor, "Multiuser detection of overloaded systems employing HARQ," in *Proc. IEEE Int'l. Conf. on Commun. Systems (ICCS)*, 21- 23 Nov. 2012, pp. 300 –304. 5, 47, 69

[22] H. Z. B. Chen, "Signal design for MIMO-OFDM systems," Ph.D. dissertation, The University of British Columbia, 2010. 5

[23] D. M. Rankin, D. P. Taylor, and P. A. Martin, "Improved information outage rate in certain MIMO systems," *IEEE Sig. Process. Lett.*, vol. 13, no. 7, pp. 393 –396, Jul. 2006. 5, 6, 11, 47, 50, 69, 102

[24] X. Liang, C. Zhao, and Z. Ding, "Sequential linear MIMO precoder optimization for Hybrid ARQ retransmission of QAM signals," *IEEE Commun. lett.*, vol. 15, no. 9, pp. 913 –915, Sep. 2011. 5, 47, 51

[25] J. N. Laneman, "Cooperative diversity in wireless networks: Algorithms and architectures," Ph.D. dissertation, Massachusetts Institute of Technology, 2002. 6

[26] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inform. Theory*, vol. 50, no. 12, pp. 3062 –3080, Dec. 2004. 6, 41, 80

[27] K. S. Gomadam and S. A. Jafar, "Duality of MIMO multiple access channel and broadcast channel with amplify-and-forward relays," *IEEE Trans. Commun.*, vol. 58, no. 1, pp. 211 –217, Jan. 2010. 6

[28] I. Chatzigeorgiou, W. Guo, I. J. Wassell, and R. A. Carrasco, "Exact and asymptotic outage probability analysis for decode-and-forward networks," *IEEE Trans. Commun.*, vol. 59, no. 2, pp. 376 – 381, Feb. 2011. 6

[29] E. C. van der Meulen, "Transmission of information in a t-terminal discrete memoryless channel," Ph.D. dissertation, University of California, 1968. 6, 38

[30] ——, "Three-terminal communication channels," *Adv. Appl. Probability*, vol. 3, no. 1, pp. 120 –154, 1971. 6, 38

[31] T. M. Cover and A. E. Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inform. Theory*, vol. 25, no. 5, pp. 572 –584, 1979. 6, 38, 43

[32] K. Ishibashi and H. Ochiai, "Analysis of instantaneous power distributions for non-regenerative and regenerative relaying signals," *IEEE Trans. Wireless Commun.*, vol. 11, no. 1, pp. 258 –265, Jan. 2012. 6, 41

[33] Y. Qi, R. Hoshyar, M. A. Imran, and R. Tafazolli, "H^2-ARQ-relaying: Spectrum and energy efficiency perspectives," *IEEE J. Select. Areas Commun.*, vol. 29, no. 8, pp. 1547–1558, Sep. 2011. 7, 8, 64

[34] B. Zhao and M. C. Valenti, "Practical relay networks: a generalization of hybrid-ARQ," *IEEE J. Select. Areas Commun.*, vol. 23, no. 1, pp. 7–18, Jan. 2005. 7, 64

[35] S. Lee, W. Su, S. Batalama, and J. D. Matyjas, "Cooperative decode-and-forward ARQ relaying: Performance analysis and power optimization," *IEEE Trans. Wireless Commun.*, vol. 9, no. 8, pp. 2632–2642, Aug. 2010. 7, 8, 64, 66, 81

[36] J. S. Harsini, F. Lahouti, M. Levorato, and M. Zorzi, "Analysis of non-cooperative and cooperative type II hybrid ARQ protocols with AMC over correlated fading channels," *IEEE Trans. Wireless Commun.*, vol. 10, no. 3, pp. 877–889, Mar. 2011. 7, 8

[37] S. R. Khosravirad, L. Szczecinski, and F. Labeau, "Rate-adaptive HARQ in relay-based cooperative transmission," in *Proc. IEEE Int'l. Conf. on Commun. (ICC)*, 9-13 Jun. 2013, pp. 5328 – 5333. 7, 8

[38] I. Krikidis, "Distributed truncated ARQ protocol for cooperative diversity networks," *IET Commun.*, vol. 1, no. 6, pp. 1212–1217, Dec. 2007. 7, 8

[39] H. Boujemaa, "Delay analysis of cooperative truncated HARQ with opportunistic relaying," *IEEE Trans. Veh. Tech.*, vol. 58, no. 9, pp. 4795–4804, Nov. 2009. 8

[40] M. B. Said and H. Boujemaa, "Throughput of cooperative harq protocols with opportunistic relaying over rayleigh channels," in *Proc. 5th IEEE Int'l. Symp. on Commun., Control and Sig. Process. (ISCCSP)*, 2- 4 May 2012, pp. 1 –5. 8

[41] B. Maham, A. Behnad, and M. Debbah, "Analysis of outage probability and throughput for half-duplex hybrid-ARQ relay channels," *IEEE Trans. Veh. Tech.*, vol. 61, no. 7, pp. 3061–3070, Sep. 2012. 8, 64, 84

[42] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, no. 8, pp. 1451 –1458, Oct. 1998. 8, 20

[43] Y. Xie, D. Gunduz, and A. Goldsmith, "Multihop MIMO relay networks with ARQ," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Nov. 30- 4 Dec. 2009, pp. 1–6. 8, 64, 65

[44] Z. Li, H. Wang, G. Wu, and T. Zhou, "Joint maximum number of transmissions handling for HARQ protocol in relay networks," in *Proc. 21st IEEE Int'l. Symp. on Personal, Indoor and Mobile Radio Commun. (PIMRC)*, 26-30 Sep. 2010, pp. 1327–1331. 8

[45] K. Zheng, L. Hu, W. Wang, and L. Huang, "Performance analysis of HARQ transmission in cooperative DF relaying systems," *Wireless Pers. Commun.*, vol. 55, pp. 441 –455, Nov. 2010. 8

[46] S. H. Kim, S. J. Lee, D. K. Sung, H. Nishiyama, and N. Kato, "Optimal rate selection scheme in a two-hop relay network adopting Chase combining HARQ in Rayleigh block-fading channels," in *Proc. IEEE Wireless Commun. and Networking Conf. (WCNC)*, 1-4 Apr. 2012, pp. 1561–1566. 8, 9, 83, 87, 91

[47] J. Choi, W. Xing, D. To, Y. Wu, and S. Xu, "On the energy efficiency of a relaying protocol with HARQ-IR and distributed cooperative beamforming," *IEEE Trans. Wireless Commun.*, vol. 12, no. 2, pp. 769–781, Feb. 2013. 8, 9, 64, 65

[48] G. Caire and D. Tuninetti, "ARQ protocols for the Gaussian collision channel," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1971–1988, Jul. 2001. 10, 88

[49] A. Zajić, *Mobile-to-Mobile Wireless Channels*. Artech House, 2013. 14, 15

[50] H.-C. Yang and M.-S. Alouini, *Order Statistics in Wireless Communications: Diversity, Adaptation, and Scheduling in MIMO and OFDM Systems*. Cambridge University Press, 2011. 15, 16

[51] K. Pahlavan and P. Krishnamurthy, *Principles of wireless access and localization*. John Wiley & Sons, 2013. 15, 16

[52] B. Sklar, *Communications Fundamentals and Applications*, 2nd ed. Prentice Hall, 2002. 16

[53] M. Ibnkahla, *Signal Processing for Mobile Communications Handbook*. CRC Press, 2005. 17

[54] C. Studer, "Iterative MIMO decoding: algorithms and VLSI implementation aspects," Ph.D. dissertation, Hartung-Gorre Publisher, 2009. 19, 25

[55] T.-D. Chiueh, P.-Y. Tsai, and I.-W. Lai, *Baseband receiver design for wireless MIMO-OFDM communications*, 2nd ed. John Wiley & Sons, 2012. 19, 20

[56] P. M. Shankar, *Fading and Shadowing in Wireless Systems*. Springer, 2012. 19, 34

[57] M. Jankiraman, *Space-Time Codes and MIMO Systems*. Artech House, 2004. 19

[58] A. Mohammadi and F. M. Ghannouchi, RF *Transceiver Design for* MIMO *Wireless Communications*. Springer, 2012. 19

[59] L. Hanzo, Y. Akhtman, L. Wang, and M. Jiang, MIMO-OFDM *for* LTE*, WiFi and WiMAX:Coherent versus Non-coherent and Cooperative Turbo Transceivers*. John Wiley & Sons, 2011. 20

[60] M. Sellathurai and S. Haykin, *Space-Time Layered Information Processing for Wireless Communications*. John Wiley & Sons, 2009. 25, 49

[61] S. Verdú, "Minimum probability of error for asynchronious gaussian multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 32, no. 10, pp. 85 –96, Jan. 1986. 25

[62] M. L. Honig, *Advances in Multiuser Detection*. John Wiley & Sons, 2009. 25

[63] P. Fertl, J. Jaldén, and G. Matz, "Performance assessment of MIMO-BICM demodulators based on mutual information," *IEEE Trans. Sig. Process.*, vol. 60, no. 3, pp. 1366 –1382, Mar. 2012. 26, 29, 30

[64] L. Sugiura, S.; Sheng Chen; Hanzo, "MIMO-aided near-capacity turbo transceivers: Taxonomy and performance versus complexity," *IEEE Commun. Surveys & Tutorials*, vol. 14, no. 2, pp. 421 –442, Second Quarter 2012. 27, 46

[65] K. Wu, L. Sang, H. Wang, C. Xiong, D. Yang, and X. Zhang, "Detection algorithm for V-BLAST systems with novel interference cancellation technique," in *Proc. IEEE, Veh. Tech. Conf. (VTC)*, Apr. 2009, pp. 1–5. 27

[66] G. Liva, S. Scalise, E. Paolini, and M. Chiani, "Turbo codes based on time-variant memory-1 convolutional codes over Fq," in *Proc. IEEE, Int'l Conf. on Commun. (ICC)*, 5-9 Jun. 2011, pp. 1–6. 30

[67] E. Rosnes, "On the minimum distance of turbo codes with quadratic permutation polynomial interleavers," *IEEE Trans. Inform. Theory*, vol. 58, no. 7, pp. 4781–4795, Jul. 2012. 30

[68] A. E. Pusane, R. Smarandache, P. O. Vontobel, and D. J. Costello, "Deriving good LDPC convolutional codes from LDPC block codes," *IEEE Trans. Inform. Theory*, vol. 57, no. 2, pp. 835–857, Feb. 2011. 30

[69] G. Han, Y. L. Guan, and X. Huang, "Check node reliability-based scheduling for BP decoding of non-binary LDPC codes," *IEEE Trans. Commun.*, vol. 61, no. 3, pp. 877–885, Mar. 2013. 30

[70] A. N. Assimi, "Diversity techniques for HARQ transmissions over frequency-selective channels," Ph.D. dissertation, University of Cergy-Pontoise, 2009. 34, 36

[71] E. Hossain, D. I. Kim, and V. K. Bhargava, *Cooperative Cellular Wireless Networks*. Cambridge University Press, 2011. 38

[72] K. J. R. Liu, A. K. Sadek, W. Su, and A. Kwasinski, *Cooperative Communications and Networking*. Cambridge University Press, 2009. 39, 42, 80

[73] H. Ding, J. Ge, D. da Costa, and Z. Jiang, "A new efficient low-complexity scheme for multi-source multi-relay cooperative networks," *IEEE Trans. Veh. Tech.*, vol. 60, no. 2, pp. 716–722, Feb. 2011. 40

[74] P. Xu, Z. Ding, X. Dai, I. Krikidis, and A. Vasilakos, "A novel relay-assisted protocol for cooperative multiple access networks," in *Proc. IEEE, Int'l Conf. on Commun. (ICC)*, 10-15 Jun. 2012, pp. 4547–4551. 40

[75] R. S. Ganesan, H. Al-Shatri, T. Weber, and A. Klein, "Cooperative zero forcing in multi-pair multi-relay networks," in *Proc. IEEE Int'l. Symp. on Personal, Indoor and Mobile Radio Commun. (PIMRC)*, 9-12 Sep. 2012, pp. 1740 –1745. 40

[76] Y.-W. P. Hong, W.-J. Huang, and C.-C. J. Kuo, *Cooperative Communications and Networking-Technology and System Design.* Springer, 2010. 40

[77] J. N. Laneman, G. W. Wornell, and D. N. C. Tse, "An efficient protocol for realizing cooperative diversity in wireless networks," in *Proc. IEEE Int'l. Symp. on Inform. Theory (ISIT)*, Jun. 2001, p. 294. 41

[78] X. Liu, X. Zhang, and D. Yang, "Outage probability analysis of multiuser amplify-and-forward relay network with the source-to-destination links," *IEEE Commun. Lett.*, vol. 15, no. 2, pp. 202–204, Feb. 2011. 43

[79] M. Dohler and Y. Li, *Copperative Communications: Hardware, Channel* & PHY. John Wiley & Sons, 2010. 44

[80] H. Alves and R. D. Souza, "Selective decode-and-forward using fixed relays and packet accumulation," *IEEE Commun. Lett.*, vol. 15, no. 7, pp. 707–709, Jul. 2011. 45

[81] M. Khafagy, A. Ismail, M.-S. Alouini, and S. Aïssa, "On the outage performance of full-duplex selective decode-and-forward relaying," *IEEE Commun. Lett.*, vol. 17, no. 6, pp. 1180–1183, Jun. 2013. 45

[82] R. Lin and P. A. Martin, "Hybrid ARQ schemes for non-orthogonal space-time block codes," in *Proc. IEEE Int'l. Symp. on Inform. Theory and its App. (ISITA)*, Dec. 2008, pp. 1475–1480. 59

[83] O. ur Rehman and N. Zivic, "Multi-layer data protection using N-Channel stop-and-wait Hybrid-ARQ in WiMAX," in *Proc. IEEE, Int'l Symp. on Comp. and Commun. (ISCC)*, 28 Jun. - 1 Jul. 2011, pp. 359–364. 64

[84] J. Dohl and G. Fettweis, "Energy aware evaluation of LTE Hybrid-ARQ and modulation/coding schemes," in *Proc. IEEE, Int'l Conf. on Commun. (ICC)*, 5-9 Jun. 2011, pp. 1–5. 64

[85] T. V. K. Chaitanya and E. G. Larsson, "Superposition modulation-based symmetric relaying with hybrid ARQ: Analysis and optmization," *IEEE Trans. Veh. Tech.*, vol. 68, no. 8, pp. 3667–3683, Oct. 2011. 66

[86] P. Wu and N. Jindal, "Performance of hybrid-ARQ in block-fading channels: A fixed outage probability analysis," *IEEE Trans. Commun.*, vol. 58, no. 4, pp. 1129–1141, Apr. 2010. 68, 83

[87] T. E. Hunter, S. Sanayei, and A. Nosratinia, "Outage analysis of coded cooperation," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 375–391, Feb. 2006. 82

[88] A. Hedayat and A. Nosratinia, "Outage and diversity of linear receivers in flat-fading MIMO channels," *IEEE Trans. Sig. Process.*, vol. 55, no. 12, pp. 5868–5873, Dec. 2007. 84

[89] Y. Jiang, M. K. Varanasi, and J. Li, "Performance analysis of ZF and MMSE equalizers for MIMO systems: An in-depth study of high SNR regime," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 2008–2026, Apr. 2011. 84, 85

[90] R. Wolff, *Stochastic Modeling and the Theory of Queues*. Prentice-Hall, 1989. 88

[91] R. Lehnert, P. Tran-Gia, and J. Charzinsky, *Providing Quality of Service in Heterogeneous Environments*. Elsevier, 2003. 88

[92] S. Marcille, P. Ciblat, and C. J. L. Martret, "A cross-layer HARQ scheme robust to imperfect feedback," in *Proc. IEEE, Asilomar Conf. on Sig., Systems and Comp. (ASILOMAR)*, 4 -7 Nov. 2012, pp. 143–147. 104

[93] J. Wang, X. Chai, X. Zhang, Z. Zhang, and K. Long, "Adaptive power allocation in cooperative OFDM transmission with three-hop DF relaying," in *Proc. IET, Int'l Conf. on Info. and Commun. Tech. (IETICT)*, 27 -29 Apr. 2013, pp. 387–393. 105

[94] P. Larsson, B. Smida, T. Koike-Akino, and V. Tarokh, "Analysis of network coded HARQ for multiple unicast flows," *IEEE Trans. Commun.*, vol. 61, no. 2, pp. 722–732, Feb. 2013. 105

[95] H. H. Kha, H. D. Tuan, and H. H. Nguyen, "Joint optimization of source power allocation and cooperative beamforming for SC-FDMA multi-user multi-relay networks," *IEEE Trans. Commun.*, vol. 61, no. 6, pp. 2248–2259, Jun. 2013. 105

[96] J. Zou and H. Xu, "Auction-based power allocation for multiuser two-way relaying networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 1, pp. 31–39, Jan. 2013. 105

[97] W. Guan, "Cooperative communication with wireless network coding," Ph.D. dissertation, University of Maryland, College Park, 2013. 105

[98] Y. Lang, D. Wübben, A. Dekorsy, V. Braun, and U. Doetsch, "Improved HARQ based on network coding and its application in LTE," in *Proc. IEEE Wireless Commun. and Networking Conf. (WCNC)*, 1-4 Apr. 2012, pp. 1958–1963. 105